ESSAYS ON ECONOMETRICS AND PLANNING

PRESENTED TO
PROFESSOR P. C. MAHALANOBIS

ON THE OCCASION OF HIS
70TH BIRTHDAY

Edited by
C. R. RAO
In Collaboration with
D. B. LAHIRI K. R. NAIR
P. PANT S. S. SHRIKHANDE

PERGAMON PRESS
OXFORD • LONDON • NEW YORK • PARIS • FRANKFURT

STATISTICAL PUBLISHING SOCIETY
CALCUTTA

PERGAMON PRESS LTD.
Headington Hill Hall, Oxford.
4 & 5 Fitzroy Square, London W.1.

PERGAMON PRESS INC.
122 East 55th Street, New York 22, N.Y.

GAUTHIER-VILLARS
55 Quai des Grands-Augustins, Paris 6.

PERGAMON PRESS G.m.b.H.
Kaiserstrasse 75, Frankfurt am Main.

STATISTICAL PUBLISHING SOCIETY
204/1 Barrackpore Trunk Road, Calcutta-35.

Distributed outside India by
PERGAMON PRESS

COPYRIGHT ©
STATISTICAL PUBLISHING SOCIETY

Library of Congress Card No. 63-21724

DLDI, NCERT
338.9
RAO 60176

Printed in India
at the Eka Press, Calcutta-35

ESSAYS ON ECONOMETRICS AND PLANNING

Contents

		Page
Foreword	*C. D. Deshmukh* ...	1
Conflict, cooperation, competition, and cupid	*Russell L. Ackoff* ...	7
Planning in Britain	*T. Balogh* ...	19
Parametric solution and programming of the Hicksian model ...	*Ragnar Frisch assisted by Ashok K. Parikh* ...	45
The national balance-sheet of the Soviet Union	*Raymond W. Goldsmith* ...	83
Elements of a theory of planning ...	*Nathan Keyfitz* ...	103
The three basic purposes of an economic system : To achieve optimum growth : To apportion production wisely : To maintain economic justice ...	*Leon H. Keyserling* ...	127
Elements of a theory of system behaviour	*Oskar Lange* ...	137
Scientific approach in planning ...	*M. Mukherjee* ...	153
A philosophical view of civilization	*Radhakamal Mukherjee* ...	177
A critical appraisal of the concept and theory of unemployment	*Gunnar Myrdal* ...	183
The role of Central Statistical Organisation in the Indian Statistical System ...	*K. R. Nair* ...	205
The strategy of Indian planning ...	*I. G. Patel* ...	217
Science as a dialectical unity of system, method, and practice ...	*T. D. Pavlov* ...	225
Measure, quality, and optimum scale	*Nicholas Georgescu-Roegen* ...	231

Page

Capital formation and its financing in the private corporate sector in India ... ... *N.S.R. Sastry and V. K. Sastry* ... 257

Planning and economic growth ... *J. Katz-Suchy* ... 267

Models for demand projections ... *Richard Stone* ... 271

India since independence : some random thoughts ... ... *S. Subramanian* ... 291

Project appraisal : a traditional approach ... ... *J. Tinbergen* ... 295

The twentieth century trend in employment in manufacture in India—as illustrated by the case of West Bengal ... *Alice and Daniel Thorner* ... 301

A letter report to Professor P. C. Mahalanobis ... ... *H. O. A. Wold* ... 309

Scientific contributions of Professor P. C. Mahalanobis ... *Editors* ... 321

Messages 343

FOREWORD

The friends and well-wishers of the Indian Statistical Institute have laid those connected with the Institute under a load of gratitude by conceiving of and executing this plan of bringing out a volume of essays to be presented to Professor P. C. Mahalanobis on his Seventieth birthday in June, 1963. To one like me who can claim to have been a close friend of Professor Mahalanobis for over 40 years and who has had the privilege of working with him in guiding the growth and development of the Institute for the last 18 years, this gesture of regard and good-will is especially gratifying, as well as appropriate, for the attainment of the three-score years and ten span by Professor Mahalanobis happens to mark the consummation of a most remarkable period of achievement in the life of this eminent scholar and scientist, restless visionary and planner.

Mahalanobis had left Cambridge, with a brilliant degree in Mathematics and Physics, by the time I was admitted there in 1915. But a close friend of his who continued his studies became a close friend of mine also, and through him I came to know Mahalanobis sufficiently well to stay as his guest in Calcutta in October, 1920, when I visited that city to appear in some Oriental Languages examination.

It was during that visit that I learnt of the high standing of the Mahalanobis family in Bengalee society, as well as getting a glimpse of Mahalanobis's interest in statistics—connected at that time with meteorology, purely as a powerful side interest of a professional teacher of physics. Mahalanobis introduced me to his aunt, a daughter of Keshab Chandra Sen, and in his company I called on Dr. Brajendra Nath Seal and Dr. Nilratan Sircar. A year later, I visited Calcutta again and got to know Srimati Nirmalkumari Maitra, then Mahalanobis's fiancée and soon after Mahalanobis's wife. Dr. Heramba Chandra Maitra, her father, was a strict Brahmo and there was much ado over the proposed mode of marriage ceremony but love triumphed in the end.

I learnt of the difficulties that had been overcome only years later, as between 1921 and 1939 I had very little contact with the Mahalanobises. The threads of friendship, not any the worse for desuetude were resumed in Simla where I came to be posted as Joint Secretary to the Government of India in the Department of Education, Health and Lands and whither Mahalanobis had come to persuade the Government of India to give him a small grant—not of more than four digits—for developing his statistical laboratory and Institute. For, in the meanwhile this was the stable but unsubstantial form his interest in statistics had taken, an interest which had grown into a passion as the years went by and the time for his retirement from his official post of Professor of Physics at the Presidency College, Calcutta drew nearer. Some space, sparing which was becoming increasingly difficult to an expanding department, had been permitted to be used for the purposes of the Statistical Institute and Laboratory.

My joining the Reserve Bank of India in 1939 as Secretary, to rise to be Deputy Governor (1941) and Governor (1943) marked a turning point in my own relations with Mahalanobis and my interest in his extra-professional work. The Reserve Bank top-level staff, with Headquarters in Bombay, spends about three months of the year (from mid-November to mid-February) in Calcutta, in order to maintain contact with that of other great industrial and commercial capital of the country. My duties therefore involved a three-month sojourn in Calcutta and far more extended opportunities of keeping up contacts with the Mahalanobises. I had taken the opportunity of this sojourn to learn a little Bengali, in order to know better my many friends in Bengal—from Cambridge days, principally. And this forged another link in my friendship with the Mahalanobises, as I discovered how closely they had been associated with Rabindranath Tagore and Santiniketan in its early days. Mahalanobis had been Secretary of Visva-Bharati and Rani (the name by which Tagore called Srimati Nirmalkumari Mahalanobis) a deeply devoted ministering angel to the Poet.

Many a Sunday I spent with the Mahalanobises, upto 1942, in a house beyond Baranagore Railway Station and after 1942 in 204 Barrackpore Trunk Road, the present somewhat bewildering Panch-Mahal like house. Indeed, I was one of those who strongly advised the purchase of that property—during the war—as I felt attracted by its two tanks and gardenland, mangoes predominating (Amrapali was the apposite name given to it by Rabindranath Tagore). The war led to many profound changes in the location and organization of work of the Institute. Some staff had to be located in far away Giridih in the Santhal country. The Presidency College became less and less of an asylum and the precincts of Amrapali took up the strain.

It was soon after my joining the Reserve Bank that I became a member of the Indian Statistical Institute, to become its President in 1944. Although I have studied lower mathematics it is not a subject I am at ease in or have a talent for, and my only connection with statistics had been carrying out random agricultural crop sample surveys, after Hubback, as a Settlement officer in Raipur in Madhya Pradesh (then Central Provinces and Berar) between 1926 and 1931, as a cross check on soil classification, correlated to soil analysis, and rent-rates fixed by me. However, increasingly I felt a sense of involvement in Mahalanobis's work in statistics and a growing desire to help him in obtaining financial support from the Government of India. As Deputy Governor and later Governor of the Reserve Bank of India, I had some influence with the Finance Ministry and felt I should use it in what appeared to me to be a decidedly deserving cause.

But statistics was not the only bond between the Mahalanobises and myself. Apart from their special relations with the Poet, I discovered Mahalanobis's deep interest in the social and intellectual movements in Bengal, especially in the life and work of Raja Ram Mohun Roy, and I learnt of his great competence as a speaker and writer of Bengali. Tagore became more and more a powerful theme of talk with us—the room reserved for him in Amrapali, from which the Poet could greet the morning sun if ever he stayed with them; the numerous tours in different parts of the world in which the Mahalanobises had accompanied the Poet, incidents of these travels that it was a pleasure to recall, particularly the triumphs; the Poet's sense of humour; his amazing speed in composing poetry, often a flow like that of the great river in a sailing boat on which he had loved to write; and the Poet's many letters to Rani.

It was inevitable that the Mahalanobises and I should make plans for me to pay my respects to the Poet, and this was hastened by his illness at Santiniketan. The Mahalanobises and I went to the Abode of Peace in February 1941, just for a day or two. I ventured to take a small offering, in the shape of a poem in Sanskrit, which the Poet received graciously and in return for which at my request he read out one of his own (*Abirbhavo*), selected by Rani Mahalanobis. Alas, that was the first and the last time that I was to see the Poet. I returned to Bombay, to my duties in the Reserve Bank, and in August 1941 the sad tidings reached me that the Poet had passed away.

I became the Governor of the Reserve Bank in February, 1943 with a more perceptible capacity to interest the Government of India in helping the Indian Statistical Institute, and in 1945 the Education Ministry agreed to make its first large and regular grant of Rupees Five Lakhs to the Institute for its Research and Training School. For Mahalanobis the period after the end of the war was one of more frequent professional trips abroad—he had in the meanwhile been made a Fellow of the Royal Society of U.K. for his work on large-scale agricultural sample surveys—and had also been elected a member of the International Statistical Institute. It was about this period that he resolved in his mind and discussed with me his idea of initiating a National Sample Survey as well as his growing interest in Statistical Quality Control.

This latter interest matured in the shape of a plan to invite Dr. Walter Shewhart, Director of the Research Laboratory of the Bell Telephone Co., the Father of Statistical Quality Control to India, in an effort to arouse interest in the movement in the business world of India. This effort had the welcome support of Dr. Lal Verman of the Indian Standards Institution and took shape in 1948. Dr. Walter Shewhart, accompanied by his wife Edna Shewhart, came to India, stayed in Calcutta with the Mahalanobises at Amrapali and in Bombay with me at the Reserve Bank House. Whilst the visit sowed the seed of statistical quality control in India, which after a late germination is at last showing signs of growing into a tree of considerable size, it laid the foundation of a deep and abiding friendship between the Shewharts and the

Mahalanobises on one hand, and the Shewharts and myself on the other, a charmed circle to which my wife, Srimati Durgabai Deshmukh, was admitted on our marriage in January, 1953. Edna is *didi* to both Rani and Durgabai, and the Shewhart home in New Jersey is our American home.

Soon after relinquishing the Governorship of the Reserve Bank in 1949, I was appointed, on the suggestion of the late Dr. John Mathai, then Union Finance Minister, Financial Ambassador in U.S.A. and Europe to the Government of India, and in that capacity accompanied Prime Minister Jawaharlal Nehru in part of his travels on his first visit to the U.S.A. in October-November 1949. Soon after that I was called to Delhi for consultations, but stayed to organize the Planning Commission for the Prime Minister and to become the Union Minister of Finance in May 1952 as a result of a concatenation of events which this is not the place or occasion to elaborate.

It is necessary to state at this stage that Mahalanobis had known Shri Jawaharlal Nehru and had had occasion to inform him of what he had been doing to promote statistics in India. Once, I remember, Shri Nehru had come to Bombay to speak on South and South-East Asia and International Relations before the Bombay branch of the Indian Council of World Affairs. That evening I was invited to dine with Shri Nehru by his sister Srimati Huthee Singh, and I had occasion to refer to Mahalanobis's work for statistics and to discover that Shri Nehru was appreciative of it and conscious of the need to encourage it. This friendly interest was to bear fruit later in the appointment of Mahalanobis as Statistical Adviser to the Government of India by Prime Minister Nehru.

My coming on the scene as Minister of Finance led to many important developments of significance to the growth of the country's statistical apparatus, developments which would not have been possible but for Prime Minister Nehru's awareness of the role of statistics in planned economic development. The institution of a Central Statistical Office, the establishment of the National Sample Survey, the regular annual compilation of the National Income Report, with intensive statistical investigations within these fields, whenever and wherever required, the formation of a Statistical Quality Control Policy Advisory Committee—all these were the product of the conjunction of the favourable circumstances which provided for the Indian Statistical Institute and Mahalanobis's ever-burgeoning ideas about the range of its potential services the strongest possible support in the Union Government, in the shape of the understanding by the Prime Minister of the importance of statistics, organized by a non-official, academically oriented and professionally competent body like the Indian Statistical Institute. Apart from the proliferation of its activities as indicated above, this period (1950–1956) saw the strengthening of the Research and Training School and its diversification, including an International Statistical Training Centre, thanks to the support extended by the Government of India.

The recognition of the Indian Statistical Institute into an Institution of National Importance by means of a piece of legislation most understandingly piloted by the Prime Minister himself, about a couple of years ago, put the coping stone on this imposing edifice the rising of which owes so much to the enthusiasm, energy, initiative and professional competence of Mahalanobis.

The last eight years have seen Mahalanobis grow into a figure of national importance. His name leapt into prominence with his putting out a draft outline of the Second Five Year Plan in the beginning of 1956. Much preliminary work, including the labours of well-known foreign specialists, had gone into this venture. That the Second Plan as finally formulated contained much of the philosophy underlying the Draft Outline testifies to the deep thinking that informed it. Since then Mahalanobis has had an assured place as a *de facto* member of India's Planning Commission, and today he is in charge both of Science and Perspective Planning as a member of the Commission—still *de facto*, because of the impossibility of divorcing him from his major interest: the Indian Statistical Institute.

As President of the National Institute of Sciences of India, Mahalanobis had recently an opportunity to contribute to the stabilization of the resources of that eminent body, including the completion of its new home on Tilak Marg, New Delhi.

It is inevitable that with his assured status in the world of science, statistics and planned economic development Mahalanobis should be frequently called upon to participate in international conferences or to advise Governments of developing countries. With no political predilections he has friends all over the globe and easy access to scientific circles in the countries of both the Western and the Eastern Blocks. Scientists and academicians from both blocks respond to his invitation to spend some time at the Institute as visiting professors or experts. Ronald Fisher, Simon Kuznets, Richard Stone, Frank Yates, Norbert Wiener, John Strachey, J. K. Galbraith, Nicholas Kaldor, A. N. Kolmogorov, J. B. S. Haldane and numerous others have at one time or another helped in the work of the Institute and added new ranges or dimensions to it. Today the Indian Statistical Institute, with its journal Sankhyā, its Ural Computer and its plans for the manufacture of calculating machines can claim to be in the forefront of scientific institutions of international importance, and its School of Research and Training, under its highly competent Director, C. R. Rao, confers graduate and post-graduate degrees of B. Stat. and M. Stat., and research degree of Ph.D. which are recognized as the hall-mark of advanced training in statistical science and practice.

With all his remarkable achievements, Mahalanobis remains a modest man. His absorption in his work and his catholicity of interest make him a man not too easy to get on with. He has no small talk and little capacity to compromise with unreason. But he has no malice and no guile—indeed the intricacies of bureaucratic administration non-plus him—it is here that I have been able to help him a little and to make his path a little easier.

Mahalanobis's eternal preoccupation with work suppresses a genuine capacity for affection—which in any case he would think it his duty to conceal—except towards canine pets. His absent-mindedness is often the butt of Rani's affectionate badinage. It is probably true to say that without Rani's unremitting help and attention, Mahalanobis would make a mess of his foreign travel, leave things behind and catch the wrong planes.

But all these are superficial shortcomings. Functionally, he is extremely efficient and the volume and despatch of his work are amazing. The proudest thing that a friend like me, of over 42 years of standing, can say is that Mahalanobis and I have never had the slightest misunderstanding, not to speak of a quarrel. May he be spared for many more years to enrich the intellectual, academic and scientific life of the country and raise ever higher the stature of her statisticians.

C. D. DESHMUKH

CONFLICT, COOPERATION, COMPETITION, AND CUPID

By RUSSELL L. ACKOFF

Case Institute of Technology

INTRODUCTION

This paper is primarily concerned with four concepts : (1) cooperation, (2) conflict, (3) exploitation, and (4) competition. My purpose is to provide operational definitions and measures of these concepts and to show how they are related. I will use these definitions and measures to classify and discuss ways of affecting conflict between individuals and groups. It is my hope that this classification will shed some light on what we can do about conflict. I also intend to raise the question as to whether modern technology can replace Cupid's bow and arrows with weapons which man can use as Cupid did.

In order to define cooperation and conflict certain other concepts are needed. In the formulation provided here the central notion is that of a "purposeful state" of an individual or group. In earlier works (Ackoff, 1957 and 1962) I have used this same starting point for developing the concepts involved in communication theory and decision theory. It should be no surprise that theories of cooperation-conflict, communication, and decision making can rest on a common conceptual foundation. This should facilitate relating and generalizing theories which deal with combinations of these phenomena.

We can build up a definition of a purposeful state as follows :

I = an individual or a group capable of displaying behaviour;

N = the environment of I; the set of things, events, and their properties that affect I's behaviour (including other individuals in some cases).

$C_i (1 \leqslant i \leqslant m)$ = courses of action available to I in N;

$O_j (1 \leqslant j \leqslant n)$ = possible outcomes of the C_i when selected by I in N.

A fuller treatment of these variables can be found in Ackoff (1957). It is convenient to define the courses of actions and outcomes so that they form exclusive and exhaustive sets. This can easily be done by taking any set of these elements and forming combinations of them by a Boolean expansion. For example, if there are two nonexclusive and/or nonexhaustive courses of action of interest, c_1 and c_2, we can form the following exclusive and exhaustive set :

$$\begin{aligned} C_1 &= c_1 \text{ and } c_2 \\ C_2 &= c_1 \text{ and } c_2' \\ C_3 &= c_1' \text{ and } c_2 \\ C_4 &= c_1' \text{ and } c_2' \end{aligned}$$

where c_i' means "not c_i." From here on I assume that we are dealing with exclusive and exhaustive sets of courses of action and outcomes.

Three measures relate the variables which have been defined:

$P_i = P(C_i | I, N)$ = the probability that the individual or group (I) will select course of action (C_i) in the environment (N); that is, his *probability of choice.* Note: $\sum_i P_i = 1.0$ since the C_i's are assumed to form an exclusive and exhaustive set.

$E_{ij} = P(O_j | C_i, I, N)$ = the probability that outcome O_j will occur if I selects C_i in N; that is, the *efficiency* of C_i for O_j in N. Note: $\sum_j E_{ij} = 1.0$.

U_j = the utility (relative value) which I places on O_j in N.

Now an individual or group (I) in an environment (N) can be said to be in a purposeful state if the following conditions hold.

(1) There are at least two courses of action, C_1 and C_2, for which $P_1 > 0$ and $P_2 > 0$; that is, I has a potential choice.

(2) The individual (I) is not indifferent to the possible outcomes of his action; that is, there is at least one outcome (O_1) whose utility is not equal to that of another outcome (O_2): there exists (O_1 and O_2) : $U_1 \neq U_2$.

(3) C_1 and C_2 have some, but unequal, efficiency for a desired outcome, O_1; that is, $E_{11} > 0$, $E_{21} > 0$, and $E_{11} \neq E_{21}$. Therefore, some choice available to I in N must be capable of "making a difference" in the desired outcome.

In ordinary English, I can be said to be in a purposeful state if (a) he prefers one possible outcome to another, and (b) he has alternative courses of action which have some, but unequal effects on the unequally preferred outcomes.[1]

The expected utility of a purposeful state (S) to an individual (I) can be defined as

$$EU(S) = \sum_i \sum_j P_i E_{ij} U_j.$$

If the sum of the U_j's is equal to a quantity a, then max $EU(S) = a$. For most utility scales the lower limit of the expected utility, of course, would be equal to zero,[2] but if a state's expected utility reaches this limit it is no longer purposeful since the conditions defining purposefulness are no longer satisfied.

COOPERATION AND CONFLICT

So much for the foundation. Now let us introduce *two* individuals or groups, I_1 and I_2. Let

$EU_1(S | I_2)$ = the expected utility of S to I_1 if I_2 is present in N; that is, if I_2 is part of N.

$EU_1(S | I_2')$ = the expected utility of S to I_1 if I_2 is *not* present in N.

We can similarly define $EU_2(S | I_1)$ and $EU_2(S | I_1')$.

---

[1] "Purpose" is treated in more detail in Churchman and Ackoff (1950).

[2] I assume the minimum value of utility is zero. If negative utilities are used, appropriate adjustments can be made in the ranges of measures specified below.

If

$$EU_1(S|I_2) > EU_1(S|I'_2)$$

then I_2 increases the expected utility of S to I_1; that is, I_2 *cooperates* with I_1. Therefore, the degree to which I_2 *cooperates* with I_1 can be defined as

$$DC_{21} = EU_1(S|I_2) - EU_1(S|I'_2).$$

This quantity has a maximum value of a and a minimum value of $-a$. If this measure is negative, then I_2 is in *conflict* with I_1. Hence, conflict and cooperation are contrary (not contradictory) relations since if $DC_{21} = 0$, I_2 has no effect on I_1.

There is nothing in the definition of conflict (or cooperation) that requires that either of the parties be conscious of the relationship or that the parties intend to conflict. For example, conflict frequently arises out of ignorance of the effect of one's action on another.

EXPLOITATION

The degree to which one individual (I_2) cooperates (or conflicts) with another (I_1) does not have to equal the degree to which I_1 cooperates (or conflicts) with I_2. Thus, two individuals may affect each other differently. This difference is the essence of *exploitation*. The degree to which I_2 exploits I_1 can be defined as

$$DX_{21} = DC_{12} - DC_{21}$$

and, correspondingly,

$$DX_{12} = DC_{21} - DC_{12}.$$

The one measure is the negative of the other. This measure can range from $-2a$ to $2a$.

Using this measure we can distinguish between three kinds of exploitation. If DC_{12} and DC_{21} are both *positive but unequal*, then the two individuals cooperate with each other unequally. The one who benefits the most can be said to be a *benevolent* exploiter. This is the type of exploitation that most colonial powers have claimed for themselves if they have admitted to exploiting their colonies at all. This type of exploitation has also been used to characterize most employer-employee relations.

If DC_{12} and DC_{21} are both *negative but unequal*, then I_1 and I_2 are in conflict with each other, but unequally. The one who suffers least is the *malevolent* exploiter; for example, one who is willing to suffer if he can make another suffer more. This is usually the case where revenge is involved. Most wars have been examples of malevolent exploitation.

Finally, if one of the parties cooperates with the other but the other is in conflict with him, we have a case of what we might facetiously call "normal" exploitation. This seems to characterize the historic relationship between slave and master.

COMPETITION

In the literature of psychology, social psychology, and sociology there is a good deal of qualitative discussion about the difference between conflict and competition. One of the more commonly cited differences is "the presence of physical force" in conflict, and its absence in competition. This difference does not seem valid to me because a prize fight is normally thought of as competition, and a street brawl is usually thought of as conflict. Although the use of force or physical contact does not seem to be essential to conflict, it can play an important role in conflict, a role that I consider below.

I make no attempt here to survey the literature on the distinction between conflict and cooperation, but I do want to cite the most suggestive discussion that I have found, that of Katz and Schanck (1937). They assert essentially that competition is conflict *according to rules*; that is, it is *constrained* conflict.

Can this distinction be made more precise ? Suppose I_1 prefers outcome O_1 to outcome O_2 but I_2 prefers O_2 to O_1, where O_1 and O_2 are exclusive outcomes (i.e., they cannot occur simultaneously). Suppose further that I_1 and I_2 are in conflict relative to these outcomes, and that they are in cooperation relative to another outcome; O_3, which they both value. For example, in a tennis match between two persons

$$O_1 = I_1 \text{ wins } (I_2 \text{ loses}).$$
$$O_2 = I_2 \text{ wins } (I_1 \text{ loses}).$$
$$O_3 = \text{recreation for both } I_1 \text{ and } I_2.$$

The more intense the conflict with respect to the first two objectives, the more efficiently the third is usually served. This, I think, is the essence of competition.

The "third" outcome may be an objective of a "third party" (for example, an audience) rather than an objective of the participants in the conflict. If this third objective is not shared by the conflicting parties the competition can be said to be *extrinsic*. If it is shared by the conflicting parties, it can be said to be *intrinsic*. Competition may be both intrinsic and extrinsic as, for example, in a tennis match before an audience.

In either intrinsic or extrinsic competition it is clearly the function of the rules to assure the effectiveness of the conflict for the third (cooperative) objective. For example, it is clear that in economic competition, conflict between rival producers is supposed to serve the interests of the consumer. Laws and regulations are enacted and enforced to make sure that this is the case. In the economic system of the United States it is illegal or immoral, for example, for two "competitors" to cooperate in exploiting the consumer; they may, however, exploit each other.

In many formulations of competitive situations the cooperative objective is not made explicit but it must be there. I feel that one of the reasons that competitive models so frequently fail to predict competitive behaviour is that they fail to encompass the essential cooperative aspect of competition; they treat competition as pure conflict.

Competition, then, involves both conflict and cooperation. More precisely, it exists if

(1) there are at least two individuals, I_1 and I_2, who prefer exclusive outcomes O_1 and O_2 respectively, in an environment N.

(2) Relative to these outcomes, I_1 and I_2 are in conflict; that is, they select courses of action which reduce the expected utility of the other (consciously or unconsciously).

(3) The conflict with respect to O_1 and O_2 is efficient for (the nonexclusive) outcome O_3 which is valued either by both the conflicting parties, or a third party. Competition, then, can be characterized by three measures :

(a) $|DC_{12}+DC_{21}|$: the absolute sum that may be called the *intensity* of the conflict. (Note that if one party is in cooperation with the other, this measure decreases).

(b) $|DC_{12}-DC_{21}|$: the absolute difference that may be called the *asymmetry* of the conflict (or cooperation).

(c) The efficiency of the conflict for the "third" objective.

At present I see no suitable way of combining these measures into one, but I feel that it will eventually be done.

WHAT CAN BE DONE ABOUT CONFLICT

Rapoport (1960 and 1961) has identified and defined three modes of conflict : *fights*, *games*, and *debates*. He characterized these modes briefly as follows:

> ...in a fight the urge is to *eliminate* the opponent;
> in a game the problem is to *outwit* the opponent;
> in the debate, the goal is to *convince* the opponent.
> (1961, p. 215, italics mine).

Rapoport argues that these three modes of conflict form a hierarchy: debate as the "best" mode, fights as the "worst", and games in between. The scale on which the hierarchy is based appears to be the following :

(1) Fights are the "worst" because they are generally[1] intended to *increase* the conflict.

(2) Games are "better" because they are generally intended to *stabilize* the conflict.

(3) Debates are "best" because they are generally intended to *decrease* the conflict.

Although the outcomes—increase, stabilization, and decrease of conflict—are clearly exhaustive, it is not apparent that the modes—fights, games, and debates—are. Consequently, I should like to go back to the definition of conflict given earlier and construct an exhaustive classification of what can be done to affect it.

---

[1] Not necessarily always.

The definition of conflict which I formulated requires the presence of the conflicting parties in the same environment; that is, where they can affect each other.[1] Therefore, the conflict may be affected if

(1) either one of the parties, both, or an "outsider" removes a participant from the environment, or

(2) either one or both of the participants removes himself from the environment.

The conflict may also be affected if

(3) the environment is modified so that the effect of the behaviour of each on the other is removed. This is equivalent to dividing the environment.

These three "modes" have the common characteristic that they separate the opponents *physically*.

If I_2 is in conflict with I_1, the definition also requires that I_2's behaviour affects the outcome of I_1's behaviour. Therefore, the conflict may be affected if

(4) I_1 changes I_2's behaviour, or

(5) I_1 changes his own behaviour.

Changes in behaviour means changing either probability of choice (P_i), or efficiency of the choice (E_{ij}), or both.

Finally, conflict may be affected without changing either behaviour or the environment, if

(6) the utilities that one or both participants place on outcomes in the environment are modified.

For example, if I_1 places no value on the outcomes affected by I_2, then I_2 cannot be in conflict with I_1.

These last three "modes" have the common characteristic that they seek to separate the opponents *psychologically*.

Now I should like to consider these six modes in more detail and examine their relationship to Rapoport's three, and to their possible effects on conflict; their tendency to increase, stabilize, or decrease it.

PHYSICAL SEPARATION OF OPPONENTS

There are three ways in which the opponents can be physically separated. In the first, one or more of the opponents can actually or effectively be removed (involuntarily) from the conflict environment. In the second, either of the parties may leave on his own volition. In the third, the environment (but not the opponent) can be manipulated so that the detrimental effects of one opponent's behaviour on the other does not occur.

---

[1] They may be in different places but still affect each other, and hence are in the same environment, in the sense in which I use this term.

In the first of these, at least one of the conflicting parties is either removed from the environment or is physically restrained or incapacitated in the environment, and is thus made incapable of affecting the other. *Force* is used to remove, restrain, or incapacitate one or more parties of the conflict. It may be applied either by participants in the conflict to each other or by one or more outsiders.

Not every use of force is a fight. A fight involves the use of force with the intent to harm by two or more parties against each other. Rapoport describes a fight as "characteristic of the sub-human world... All acts are more or less determined by preceeding claims of events, not by rational choice" (1961, p. 210). A fight generally increases (intensifies) the conflict, extends its duration (by making enemies of the participants), and it may lead to the destruction of one or more of the opponents and many others who are not involved in the conflict.

Force may also be used with the intent to prevent harm, as for example, in separating parties engaged in a fight or in preventing an assault or attack by one person on another.

It should be noted, as Rapoport has done, that a threat to use force or to exact some other type of penalty which is intended to inhibit conflicting behaviour, is not the same as a fight.

"To *induce* an action...is mostly physically impossible. The most you can do is to offer a choice between alternatives, for example, "Sign this or die." We call such an offer intimidation by the use of force, but in the last analysis it is the Other who makes the choice. If he chooses not to sign, he cannot be forced to do so, because his nervous system and his muscles cannot be controlled by another in coordinated fashion (1961, p. 215)."

The objective of a threat is to change a participant's probabilities of choice (one of his purposeful-state variables) and hence involves psychological manipulation.

The second way of separating opponents involves a voluntary exodus by one or both of the parties. This may increase as well as decrease the conflict as, for example, when one of the parties desires conflict with the other. Because voluntary exodus usually involves a sacrifice by the departing party, it is not frequently used, but it can be a simple way of dissolving a conflict.

The third way of separating opponents consists of dividing the environment into two parts so that the opponents no longer affect each other. For example, if I_1 is flashing a light that distracts and annoys I_2, a screen may be used to block the light. Most conflicts arising from intrusions on the senses can be removed by modifying the environment. Here the effects of behaviour are modified, but neither the behaviour nor the parties are changed. Such a procedure is intended to weaken, if not *dissolve* the conflict. Most conflicts that arise out of scarcity can also be dissolved by modifying the environment by making available more of that which is scarce. This, of course, is not always possible or practical.

Separation of opponents by modifying the environment is an attractive way of handling conflict because it does not involve "changing people". Consent or cooperation of the opponents may not be required or may be relatively easy to obtain.

STABILIZING CONFLICT

To seek a course of action or courses of action (a strategy) which (in some sense) minimizes the detrimental effect of one opponent's behaviour on the other's expected utilities is to attempt to stabilize a conflict. The conflict can be said to be *solved* if, when either party changes his behaviour without the other doing so, the one who changes incurs an additional loss. Thus a solution is a state of equilibrium. Attempts to solve a conflict in this sense correspond to what Rapoport has called a game. So-called "rational behaviour" is sought, strategies which satisfy one of any number of criteria which game theorists have suggested (e.g., minimax, or minimax regret).

The solution to a conflict does not *eliminate* it although it may accelerate its completion; that is, its reaching an end-state, if such a state is attainable.

Those involved in a conflict which is embedded in competition are usually constrained by the (explicit or implicit) "rules of the game" either to try to solve (and hence stabilize) or to increase (but not decrease) the conflict. The rules usually specify the conditions under which an end-state (if permitted) is reached and all the conflicting parties can do, is to seek a preferred pay-off; that is, try to solve the conflict. Infraction of the rules generally involves penalties; hence, the rules constitute a threat. Where a third party is involved—for example, the audience at an athletic event—it sometimes tries to encourage an increase in conflict by offering rewards to the one that carries out the conflict most effectively (i.e., wins).

A conflict is frequently deliberately imbedded in competition in order to bring it under control and prevent great losses to either party. For example, I can remember being in a fight with another boy at age 13 in a school hall-way when the instructor in physical education stopped us and took us to the gymnasium, equipped us with heavy gloves and put us into a ring before an audience. Our anger was quickly dissipated.

In international affairs many have hoped to create a world organization that would be capable of imbedding all conflicts between nations in competition so that such conflicts could be made to serve rather than harm man. That the United Nations does not yet have the power to so modify conflicts between nations is apparent.

When two parties communicate for the purpose of resolving a conflict they are said to be *negotiating*. The way in which such negotiation is organized and the way that its environment is structured can have a considerable influence on its chances of success. Even such things as the arrangements of the room in which negotiations take place can influence the outcome. There is need for considerably more research on these aspects of negotiation.

In most collective bargaining, it seems to me, the conflicting parties are more concerned with removing exploitation (i.e., equalizing conflict) than with removing *it*. The parties are primarily motivated by the desire "not to be taken advantage of." The objective of most such efforts, therefore, is a state of conflict-equilibrium. This state, of course, may be very temporary.

The role of the arbitrator in pure conflict may be that of accelerating the formulation of a state of conflict-equilibrium. If, however, he represents a third party who benefits in a competitively-imbedded conflict, then he usually seeks an equilibrium which also effectively serves the cooperative objective in the competition.

Rapoport and others have made the important point that many conflict situations *cannot* be represented by a game-theoretic model which has a solution. One might argue that richer models are needed to make solutions possible or, as Rapoport has, that the outlook (purposeful-state properties) of the other must be changed.

"As soon as the number of players exceeds two...coalitions become possible... The problems of actual coalition formation were explicitly left out by von Neumann. These problems were taken up by others, and their investigations showed very soon that in any theory of bargaining extra-game theoretical considerations must be taken into account : Such matters as trust and suspicion; awareness of collective interest; susceptibility or immunity to threats, etc...

"We see, then, that a rigorous examination of game-like conflict leads inevitably to considerations of conflicts of still another sort — conflicts of conviction. For "what is fair" cannot be discovered by empirical evidence, nor decided by the outcome of a game but is a matter of conviction... . Usually a person acts in a socially acceptable manner only if he is convinced that others too, normally act in this way. Therefore, to make an individual social-minded instead of egocentric, we must convince him that others are social minded too (1961, pp.214-215)."

If these properties must be and can be changed to make a solution to the conflict possible, then they can also be changed to make a resolution of the conflict possible.

In order to solve any game-like conflict each player must know the other's pay off matrix. This requires understanding of the opponent's motives : why he does what he does. This understanding is precisely what is missing in most conflict situations. In order to obtain such understanding, effective communication between the opponents is usually required. But if such communication is possible, then there is a possibility of decreasing, rather than stabilizing or increasing the conflict. Put another way, no more understanding of an opponent is required to dissolve a conflict than is required to solve it.

Psychological separation of opponents

A purposeful state, it will be recalled, is characterized by three types of variables : the probabilities of choice of each course of action (P_i), the efficiency of these courses of action relative to the outcomes (E_{ij}), and the utility of the possible outcomes (U_j).

The manipulation of these variables is normally accomplished by communication: the transmission of messages between the opponents or to them from outsiders. I take such efforts to correspond with what Rapoport has called a *debate,* but I feel that this term connotes too restricted a class of communication.

As I indicated earlier, I have used the concept of a purposeful state to define concepts and measures of human communication (Ackoff, 1958). Summarizing this earlier work very briefly, messages which affect the probabilities of choice of an individual (P_i) transmit *information* to him; those which affect the efficiency of his actions (E_{ij}) transmit *instruction*; and those which affect the utility he places on outcomes (U_j) transmit *motivation*. (The amounts of these which are transmitted, in my scheme, may be negative.) A single message may have any one, two, or all three communication functions.

Communication, however, is not the only way of affecting an individual's purposeful state variables; they may also be affected indirectly by manipulating the environment. Manipulation of the environment for this purpose is different from manipulating the environment so as to separate the opponents. In the former the objective is to change behaviour; in the latter it is to change the effect of behaviour. Suppose two individuals are in an environment which contains two automobiles, one of which is hidden, and that both want to use an automobile at the same time. This conflict may be dissolved either by informing them of the availability of the second automobile or by changing the environment so that one of them "discovers" the second automobile. The purpose of this "intrusion" is to change one of the party's probability of using the second automobile; it may be accomplished with or without changing the environment.

As mentioned earlier, if only one automobile is present in the environment, the conflict may be dissolved by bringing a second automobile into the environment and thereby "separating" the opponents.

Now suppose two persons want to use the one desk calculator that is available in an environment to obtain the squares of different sets of three digit numbers. Both have reason to want to get it done as quickly as possible. If both try to enter their numbers in the same columns of the keyboard at the same time their efficiency in using the calculator is very low. They can be told, however, that if they use different columns separated by four unused columns, they can get squares of different numbers simultaneously. Such a message would *instruct* them in the use of the calculator; that is, enable them to use it more efficiently under the circumstances. It might also be possible to modify the keyboard of the calculator so that they learn how to use it simultaneously.

Finally, if two persons in an environment want things which they cannot both have at the same time (e.g., two children who want to play by themselves with the same ball), this conflict can be decreased or dissolved by reducing the value that one of the party places on his objective (frequently, by making him want something else more). This may be done either by communication (of the motivational type) or by changing the environment; for example, making available an object that is more desirable to one of the parties but not the other.

Promises or threats are types of communication which are intended to change the behaviour of people in actual (or potential) conflict. But environments can also be manipulated so that either the consequences of conflicting behaviour are made less desirable or the consequences of cooperative behaviour are made more desirable. The intent in such cases is to induce a certain kind of behaviour in the same way as a promise or threat is intended to. For example, it has been suggested that the governments of the U.S.S.R. and the U.S.A. exchange locations. The idea behind this suggestion is that if either government initiated a nuclear attack against the other country it would destroy itself. It is argued that such an implicit threat would inhibit aggressive action.

Environments can also be manipulated to provide those involved in conflict with a positive incentive to cooperation; for example, prizes or awards can be offered.

Cupid

After these sober reflections I hope you will indulge a flight of fancy and, perhaps, some wishful thinking.

The types of psychological manipulation which I have discussed in connection with conflict resolution are intended to induce nonconflicting behaviour. Rapoport—it will be recalled from an earlier quotation—observed that it is virtually impossible to induce desired behaviour by physical force. Unfortunately, this is true, with an undesirable consequence; whereas we have all kinds of physical instruments (from pins to penultimate bombs) for intensifying conflict, we have no such devices for dissolving conflict. Those who are "for peace" must direct their efforts "against war." Therefore, *pacifism* is too frequently converted to *passivism*.

When the Greeks created Cupid, a "god" who could make people love each other, they were wise enough to equip him with bow and arrows with which he could inflict love even on unwilling parties. Real (as opposed to mythical) instruments of this type have yet to be developed.

Suppose for a moment, however, that they did exist; that is, "weapons" which would produce a cooperative attitude in *A* toward *B*. If *A* "shoots" *B* with cooperation, then whatever "retaliation" *B* might seek would be desirable. The entire "logic of war and peace" could be changed if such instruments were available. Are they impossible, inconceivable, or merely undeveloped?

I do not know the answer, but certainly a part of the approximately 50 billion dollars being spent each year throughout the world on developing weapons of conflict should be devoted to exploring weapons (and other means) which will dissolve conflict. It seems to me that the physical and biological sciences (as well as the behavioral sciences) must turn at least some of their efforts to peace. They have been much too cooperative with conflict.

I know that the prospect of peace pills, peace gases, or peace bugs seems ridiculous. But how ridiculous the development of weapons capable of the complete destruction of mankind must have seemed to the ancients.

I also know that even a weapon of peace could be used in conflict. For example, one person could immobilize another with love and then slay him. But at least such "weapons" would present alternatives we do not have today. Imagine a conflict-decontamination chamber for negotiators of national conflicts. Is this any harder to imagine than was a space-chamber a few decades ago ?

This, as I have said, may be ridiculous speculation, but I cannot help but feel that a society that tries to find some positive and active (physical, psychological or social) instruments for producing peace will succeed in finding more effective ways of controlling conflict.

REFERENCES

ACKOFF, R. L. (1958): Towards a behavioral theory of communication. *Management Sciences*, **4**, 218–234.

CHURCHMAN, C. W. and ACKOFF, R. L., (1950): Purposive behaviour and cybernetics. *Social Forces*, **29**, 32–39.

KATZ, DANIEL and SCHANCK, R. L. (1938): *Social Psychology*, John Wiley and Sons, New York.

RAPOPORT, ANATOL (1960): *Fights, Games, and Debates*, The University of Michigan Press, Ann Arbor.

——— (1961): Three modes of conflicts. *Management Sciences*, **7**, 210–218.

PLANNING IN BRITAIN

By T. BALOGH

Fellow of Balliol College, Oxford

INTRODUCTION

In selecting Planning in Britain as my theme, I was led by the conviction that it was Professor Mahalanobis's most notable feat, among so many notable achievements, to have, already during the war, converted the political leaders-to-be of liberated India to the concept of deliberate planning in a democratic society. The rise of the Planning Commission is, in the main, his doing. It was he who has perceived that, if India was to emerge from the mediaeval social and primeval economic constriction, a new organ had to be created, superimposed upon, and coordinated with, the traditional administrative structure. That structure (being Victorian English in origin, and remaining so to this day in its essential character) was, and remains, quite unfit to preside over the positive task of effecting a fundamental, if peaceful, social and economic revolution in India.

He was equally influential in guiding Indian Planning away from complacent restrictiveness towards ambitious all-out effort. I was, at times, sceptical about the possibilities of carrying into effect these targets, and anxious about the effects of an attempt to reach them without the required willingness and determination for the sacrifice involved. I never doubted their absolute necessity and Professor Mahalanobis's underlying concept of India's future.

Nothing can bring out the full magnitude of Professor Mahalanobis's achievement than a discussion of the current status of economic planning in Britain. That this country is some twenty years behind India is to an appreciable extent due to his success in carrying with him the Prime Minister and other political leaders. Yet can there be any doubt about the strength in India of the administrative-classic-amateur Mandarin class. All lovers of India must fervently hope that through the turmoil of foreign politics and considerations of defence, the vision of India's leaders will remain firmly fixed on the unalterable truth that strength and security can only be achieved through successful industrial and agricultural development. All steps that endanger this will—despite short term plausibility—in the end redound to the advantage of India's enemies.

1. THE DRIFT TOWARDS PLANNING

1.1. The concept of planning has provoked bitter acrimony and suffered startling changes in fortune in Britain since the end of the war, without the more influential protagonists understanding much of the implications either for or against.

1.2. Planning was first hailed as the certain way to a new and better world in peace as it had enabled the full mobilisation of available resources in war.[1] No longer would the destiny of the nation be dominated by the uncontrolled sway of the market. Unemployment would be done away with and social priorities enforced. A new era of economic equality, stability and progress would dawn. In reality, war-time planning was rather rudimentary, crude, and negative. It was forced on a reluctant bureaucracy by the irresistible needs of defence. The Labour Government was eager to claim credit for deliberate planning, but did nothing to adapt direct controls to the new peace-time needs, and co-ordinate their work flexibly with financial incentives and penalties.

---

[1] Sir Oliver Franks, *Central Planning and Control in War and Peace*, London. Also the bombastic claims of that least fortunate among the unfortunate British Chancellors of the Exchequer, Dr. (later Lord) Dalton (*The Fateful Years*, London 1957).

1.3. In consequence one unforeseen crisis followed another. Each was unnecessary and easily preventable with some professional knowledge: coal and dollars in 1947, devaluation in 1949, rearmament in 1951, All bear testimony to the bankruptcy of the system. Dalton's memoirs give a vivid picture of the growing physical and intellectual incapacity of Labour economic policy-making after 1947, once the safety of war-time controls had been abandoned and the American Loan exhausted. Their so-called chief planner was a dilettante business man who did not believe in the task entrusted to him.

1.4. The Labour Ministers, who could not admit their abject failure to plan, in any sensible sense of that maligned word, took refuge in claiming that their failure was due to their reluctance to coerce rather than their incapacity to discern the problems. Once they had entered on the path of liberalising the economy instead of creating positive peace-time controls, they were irretrievably lost. They could not compete with the Tories in freedom, nor had they the support and confidence of the owning and managerial classes. Thus each step towards financial liberty brought nearer the point where private financial power could reimpose veto on economic policy. Planning as a concept was completely discredited and disgraced. It became a dirty word.

1.5. Next reliance on the price mechanism was saluted, especially by the Civil Service, as the one final solution to our problems. Full convertibility of a floating pound would be supplemented by monetary policy controlling the volume of currency and ensuring by its subtle influence a perfect balance between global supply and demand. The commands of the sovereign consumer would then be implemented by being transformed into profit opportunities, and incentives to higher productivity, higher production and higher investment recreated by abolishing state interference. Provided monetary balance was enforced in this indirect and impersonal way, optimum progress would automatically be within our grasp. Planning would be done by individuals and firms, co-ordinated by market forces.

1.6. Alas for these illusions; Tory "freedom" did not work better than Labour "planning." In the renewed crises, 1955–1957–1959–1961; in the purposeless alternation between policies trying to stimulate and repress demand, enforced by the monotonously recurrent deficit in the balance of payments, and by fears of crises of confidence, stultification was complete. Investment was cut and economic competitive power and growth reduced below that of our great industrial rivals. Our growing impotence was due mainly to the successful election tactics of the Conservative Cabinet which necessitated the premature release for consumption of the vast resources, some £1500m. per annum, suddenly presented to Britain by the improvement of the terms of trade. Had they been reserved for investment we should now be well on the way to catching up with the growth of Germany, if not Japan. The fitful pattern of economic policy forced on us by foreign exchange crises has cost the country an appreciable loss of income through decreased investment, productivity and growth —by now certainly not less and possibly much more than ten percent of the national income, or say £2500m. per annum. This shameful failure has resulted in British national income per head dipping below that of Germany and France for the first time in several hundreds of years. Even more ominous, Soviet output per head has now surpassed ours. Twenty years ago it was less than half. Forty years ago less than a quarter. Here is a portent which is steadfastly ignored by our leaders.

1.7. The monotonous recurrence of crises seems at last to have made some impression. Having experimented with, and failed, both in beating the trade unions into acquiescence by deflation, and in increasing the share of profits and of investment to strengthen our power of resistance, the Conservative Government seemed at the turn of 1960–61 have realised that new tactics were needed. Mr. Selwyn Lloyd, then Chancellor of the Exchequer, confronted with the 1960 crisis went somewhat beyond the conventional reaction. No doubt he followed his Labour and Conservative predecessors in adopting for the sixth time since the war an orthodox programme of restrictions. But advancing beyond them he announced that he was determined to get a new policy launched—forward looking and expansionist.

With the benefit of hindsight this all now looks ironical: Mr. Selwyn Lloyd was brusquely dismissed by Mr. Macmillan in 1962 precisely because he was in fact unable to initiate a new policy. Still he left an awkward problem for his successor because, spurred by the FBI, he committed himself more and more to some sort of deliberate planning of our economic development. It was to be Tory planning, of course, with a difference. The Japanese and, conspicuously, French planning was invoked to show that informal, persuasive planning in an unchanged economic and social framework could work successfully its miracles of refreshed dynamism. It could, it was asserted, for instance, overcome the comprehensible reluctance of single entrepreneur to chance their luck and invest at a higher rate than customary for fear that they would be left in the lurch by their fellows.

1.8. In both France and Japan the majority of industrial leaders was made to believe that the high rate of expansion was both imperative and possible. This was achieved in both countries by preparing long-term—ten to twenty year—plans of accelerated development to prove that a logical and consistent basis can be found for an upward surge, and that the entrepreneur did not have to fear failure. In both cases almost complete success crowned these efforts. Why not imitate them?

1.9. The new National Economic Development Council is the result of this belated doubt about the perfection of "free" markets, and of untrammelled working of the price-mechanisms.

1.10. If one wishes to evaluate the appropriateness of a new organisation from the view point of its declared functions, if one wishes to differentiate between an honest attempt to deal with a genuine problem or a mere deceptive covering for unadmitted ulterior designs, one must ask oneself how precisely its duty is to be implemented. Here the contrast of the British scheme with French and Japanese plans could not be more blatant. Not merely were the general economic conditions far more auspicious for success in those countries, but the arrangements were far more aptly fitted for dealing with the problems.

1.11. In both cases the planning staff was organised outside the Treasury, but with access direct to the Cabinet through a Finance Minister (who was also responsible for the National Economy) in France, and through the Prime Minister himself in Japan. They formed part of the Civil Service. Moreover, in both cases the strategic position of the Government had great advantages as compared with Britain. The high command of the Japanese industry is once more highly concentrated, following the failure of the American Occupation Authorities to smash the Zaibatsu, the great family trusts. In France, too, there is an informal control over the economy by the high bureaucracy who traditionally provide also a large part of the commanding personnel of the key institutions and firms, especially the great banks and the great insurance companies (all in national ownership). Thus a pervasive influence over industry is secured through the control over short and long-term credit. By the close relations of these new, and well-trained professional bureaucrats in and out of the Civil Service, whose technical economic competence facilitates the smooth implementation of any policy. All this contrasts regrettably with the dilettantism of our own Civil Service and the wide diffusion of power over the economy. It is obvious that the nomination of five or six industrialists with no experience in planning and no particular weight in industry, which in fact happened in Britain,[2] cannot be a substitute for this intense concentration of professional competence and close co-operation in both France and Japan between industrial leaders and Government officials.

1.12. The fact is the more significant as there are vital differences between our situation and that of the two other countries. In both these countries there was a powerful basic conjuncture of economic forces making for expansion. In Japan, devaluation aggressively cut real wages. This sharpened Japan's competitive power, and provided a positive stimulant to economic expansion. There was a huge reserve army, in agriculture and in the primitive artisan sector the weight of which kept wages relatively low, despite spectacular increases in productivity. This enabled the reconquest of foreign markets, the key to Japanese expansion: with the expansion of national income import requirements rose, but could easily be met by the sharp rise in exports.

1.13. In France economic growth during the Fourth Republic, that is until the take-over by the General in 1958, was the concomitant, if not the consequence, of the sharp inflationary increase in demand due to colonial wars and uninterrupted internal investment. The possible external effects, the tendency for imports to rise when demand rises at home, was repressed by a tight tariff policy and an almost comprehensive quota system. The French economy is, in any case, rather self-sufficient; it depends far less on imports than that of England. This relative self-sufficiency was consciously fostered in the post-war period by a system of imperial preference prevailing in the Colonies, through duties, quotas, and exchange regulations. In comparison to this the Sterling Area, even at its most restrictive, was a free trade system.

---

[2] Cf. paragraph 1.17.

1.14. When some of these buttresses were dismantled (—the overseas franc zone still remains an almost exclusive French preserve—) on joining the Common Market, a sharp devaluation of the franc reduced real international costs. Wages did not rise parallel to the rise in prices. The Trade Unions had been weakened by the failure of the general strike against the General. Thus the balance of payments abruptly improved. An excess of exports restarted domestic expansion which had been stifled at first by the orthodox deflationary policy of M. Rueff. Profits further increased, and with it investment. Finally consumption also rose. The second French economic "miracle" has been the outcome of a sharp redistribution of the national income towards profits, combined with a return of "confidence" of the entrepreneurial class as a result of the victory of the General. Eventually, *ex-post*, a rapid increase in wages became possible without inflationary effects, and as a result of the combined effect on productivity of accelerated investment and expansion. Consequently the militancy of the Trade Unions did not revive, even after the charismatic glamour of the General wore rather thin.

1.15. Thus both France and Japan planning could restrict itself to controlling and dissuading over-optimistic firms, and encouraging certain basic industries whose huge needs for capital investment make them particularly sensitive to uncertainty. For both purposes the planners possessed ample power and influence.

1.16. In Britain, at present, objective economic conditions are particularly difficult for planning expansion, and the failure of the Brussels talks does not relieve the position.[3] In view of the magnitude of the task, the new planning agency seems ill-conceived. It consists of representatives of the Government and of the great vested interests of the economy: of the industrialists and the heads of two nationalised industry; and, after some hesitation, and some qualifying conditions, of representatives of the Trade Unions. The non-Government members serve in their individual capacity. There are also two independent members, and, as the sole representative of the secretariat, its Director-General Sir Robert Shone. None of these had any special experience in, and so far as one knows would not claim any special aptitude for, large-scale economic planning. The Council is fortunate in having Sir Donald MacDougall to act as the Director of their Economic Research Division (though he is not a member of the Council itself). A talented but very small staff was recruited with great difficulties. At first there were no consequential changes even in the civil service structure itself. Subsequently, at the end of July 1962, Sir Frank Lee's retirement presented an opportunity for some overdue reform.

---

[3] The entry of Britain into the Common Market would certainly have demanded a sharp devaluation of the pound.

1.17. The British Civil Service structure evolved in the period of her economic supremacy in a framework of laisser faire. Administration had been strictly separated from technical knowledge. Its structure was based on the claim that expertise could always be secured. Pure intelligence and detachment was made the basic requirement for recruitment and for further successful promotion.

1.18. The one justification for the dilettante approach is when it works. But, who, looking at the run of fiascos since the war, can permit the Treasury that claim? There was the premature convertibility of sterling, leading to the purposeless exhaustion of the American loan and thence to devaluation. There was, to take another example, the attempt to 'let the pound find its own level' in a 'free' market, which led to the crisis of 1955 to 1957. What the Treasury officials never seemed to grasp was the fact that the restrictions necessary to maintain 'freedom' (that is, freedom of the financial interests sponsored by the Bank of England) were bound to weaken the competitive position of the economy. Each time investment was cut, demand fell and imports dropped off; and productivity also suffered. The pound could be saved, but only at the price of weakening Britain against her competitors and, once the immediate pressure was relieved, setting the scene for yet another crisis. A great deal of energy, moreover, had to be spent on exhorting the trade unions to accept a wage-freeze at a time when real wages in other countries of Europe were rising faster than ever before. There never was a sensible policy for increasing investment and productivity; there never was a social policy which would have won the active co-operation of the unions.

1.19. There were changes in the government; chancellors came and went. But nothing seemed to make difference in the Treasury at George Street—either to the official hierarchy's habits of mind or to the internal organization of the Treasury. Indeed, even the modest progress made during the war years was progressively lost. During the war there had been created an impressive Economic Selection of the Cabinet which had access to all ministers, and it survived the exodus of the temporary wartime staff which had been drawn from the universities to revolutionize British economic policy-making. If this section had survived it might, with luck, have counter-balanced the baneful influence of the Treasury and the Bank of England. But the chance was lost when Sir Stafford Cripps—who had been made economic co-ordinator after the 1947 convertibility crisis—had to replace Dr. Dalton as chancellor. He organized a so-called 'planning staff' in the Treasury, distinct from the Economic Section—and the latter was further weakened when the Treasury absorbed it. Things were not improved by Mr. Macmillan's reorganization of the Treasury: the introduction of Sir Roger Makins from the Foreign Office scarcely raised the dust in George Street, though it was humiliating to the high officials, and even that breath of wind soon passed.

1.20. In 1962 for the first time, there was a change that seems to be more than the routine shuffle caused by retirements and promotions. It is the most radical reform attempted since Lloyd George so unwisely gave the head of the Treasury supreme authority by making him also head of the whole civil service—and, later, of the secretariat of the cabinet. The job, at long last, has been broken up and divided among three people. But that is not all. There has been the sudden and imaginative elevation of relatively junior officials. (The most important post went to a man who is not yet a knight—and unprecedented shock to those with a sense of the settled decencies!) Their promotion brings to an end of the epoch, beginning with Warren Fisher and ending with Lord Bridges, in which progress in the hierarchy bore little relation to technical knowledge, or intuitive acumen and performance.

1.21. Under the new arrangements the head of the civil service will concentrate on management and personnel problems. The secretary of the cabinet will apparently be confined to that function. The successor to Sir Frank Lee at the Treasury will control national expenditure as well as financial and economic policy —a job hitherto done by Sir Norman Brook in addition to his general duties for the civil service and the cabinet. The Treasury itself is being reorganized into these three branches; the economic experts are to be integrated into the division dealing with economic policy in general (in contrast to financial and monetary policy and interdepartmental co-ordination). Such changes suggest that the day of the unchallenged domination of the dilettante may at last be coming to an end.

1.22. The division of labour at the Treasury is still not logical. Financial and monetary policy cannot be separated from general economic policy, or dealt with by men who are not economic experts, leaving co-ordination to the man at the top. The system is even less likely to work effectively if that man is also burdened with the supervision of financial control—the Treasury check on spending by other departments. For this is a task which requires technical knowledge of departmental expenditure and it ought (as in most other important countries) to be assigned to a special Ministry of the Budget. The Treasury should be able to consider the economic repercussions of the budget and it should not be encumbered with the detailed scrutiny of all expenditure. The two tasks are distinct and require a different approach.

1.23. The 'integration' of the economists with the general administrators, moreover, can be successful only if the economists really are integrated into the service: that is, accepted into the establishment and given appropriate ranks and prospects of promotion. Otherwise 'integration' may mean little more than a perpetuation of their present low status in a new guise. For the need is to spread professional knowledge through the whole civil service, not to confine it to specialized groups that are isolated and ineffectual. This, in the long run, means a change in the patterns of power and in those who wield it—and there is no guarantee that the present reforms will achieve this.

1.24. Moreover, there are no economists in the various departments who could draw up detailed suggestions for alternatives from which the central organ can make its choice to fit in with general conceptions. As the new planning mechanism is outside the normal Civil Service hierarchy, it cannot demand authoritative information or the elaboration of investment projections. Its functions in the process of framing budgetary, fiscal and other policies remain obscure. Thus the most awkward problems might arise. Either the NEDC Secretariat must duplicate the function of the economic departments of the State in order to harmonise general policies with its own plans, and it must force the Government to adopt the latter, or it has to modify its plans according to the policies of the Government. In the latter case long term planning would go by the board; in the former an ad hoc body would usurp the functions of the Government. In either case violent conflicts might arise. Whatever this arrangement and conception might be, it cannot possibly be termed long term planning. So long as a do-nothing Cabinet is combined with an ignorantly dilettante bureaucracy, progressives in this country might derive a sardonic satisfaction that poor Mr. Selwyn Lloyd (who was so cruelly and unjustly dismissed from carrying out his master's policies), should have left a thorn in the flesh of the Government and the bureaucracy in the shape of the new planning organization. But it is not a set-up which can plan in any real sense of the word.

1.25. In contrast to the conceptual and organisational weaknesses, the problems facing Britain are exceedingly complex and perplexing. Unlike France before 1958, Britain cannot now permit much increase in demand. If she did she would once more slither into a balance of payments crisis. It is doubtful even whether at the moment we are paying our way. The protective buttresses which enabled France to achieve a high investment level cannot be reproduced without far-reaching reforms in Britain. We lack the control mechanism internally externally. General de Gualle's brutal rupture of the talks with the Common Market countries about our unconditional entry has given Britain an uncovenanted new chance of putting her house into order before it joins a community consisting of more dynamic and far more competitive units, and dedicated (at least in principle) to rigid obeisance to discipline of free markets. But this new opportunity has yet to be exploited.

1.26. Nor can we benefit, as the Japanese did at the beginning of their expansion, by a tremendous filip through devaluation coupled with pressure on real wages, a situation which has been reproduced, as we have seen, in the second French miracle. A mere exhortation to industrialists will not produce the investment, the technical leap forward, coupled with the squeezing of manpower from inefficient firms in order to enable the technically up-to-date to expand. Never did Britain suffer more from the hoarding of manpower, from economic constipation than at present.

1.27. It is clear that much more is needed than the assembly of scattered investment projects. But the Government does not seem to have created a mechanism which could elaborate, and see the implementation of, a set of coherently chosen targets enforcing social priorities in which the effects of the totality of investment projects have been coherently worked out and care taken, through general economic policy measures, to make them consistent with one another, and see to it that they did not conflict with the balance in the balance of payments, yet achieve an acceleration of growth. A real development plan, moreover, must be based on a definite conception of economic advance, the projects so chosen that they implement that basic concept while at the same time forming a good basis for a coherent and self-consistent expansion in which each of the projects is economically justified.

1.28. The contrast between the magnitude of the task and the weakness of the institutional arrangements is so striking as to suggest that the Conservative Government in Britain was not so much convinced that a new national agreement about a well-conceived plan could be used to energise the country and secure new harmony for greater effort, but rather that they had hoped to use the plan as it emerged as a vote-catching gimmick for the next election, showing a rosy future for all.

1.29. There were other possible advantages too. The creation of the National Economic Development Council could serve—or so Mr. Lloyd seemed to think—a hardly less important and favourable purpose. The Trade Unions, by their maladroitness, have relieved the Tory Government of much unpopularity since 1952. It has not been the Government that got the blame for its helpless incompetence in matters economic, exemplified by our relative eclipse in the world economy, or the foolishness of the handling of our relations with the Continent. It is the Trade Unions who were blamed for restrictive practices, for strikes and high prices, and secured votes among the working-class, especially the womenfolk, for the Conservatives necessary to put the Government twice back into power. The National Economic Development Council, it was thought, might be used to perpetuate this happy state of affairs.

1.30. The Trade Unions did not dare to state clearly their price for joining the Council. Nor have they ever seriously considered (far from having reached agreement on) the conditions under which they would consider acceptance of a coherent wages policy. They are therefore in a weak position. Once they consented to participate in the work of the Council they could be forced further on to the defensive. If they showed themselves reluctant to name a price for their acquiescence in creating the conditions for stability, in agreeing to what must appear to the man in the street a nationally essential wage restraint, to harmonise the increase in wages with an increase in productivity, they might easily discredit themselves and the Labour Party. The average voter by now must be rather muddled about the rights and wrongs of the case. He, as consumer, certainly has benefited from the relative stability in prices. The NEDC, it is clear, had powerful immediate attractions for the Tory Government in 1961.

1.31. If on top of everything they would have been able to get Britain to join the Common Market by the end of 1963, the stage would have been set for a fourth electoral victory and the disintegration of the Labour Party and trade union resistance, in one word or an economic "miracle" on the French pattern of 1958. If the original tactical plan of the Conservatives had succeeded it would have been the Labour Party that would have had to defend its dissatisfaction with the entry into the Common Market as *an accomplished fact.* It would have had to propose a new change and could have been accused of a lack of internationalism, of failure to look forward. The possible unfavourable economic repercussions of joining the Common Market would not have been apparent at that stage. The tactical plan of the Conservative Government was, as always, ingenious.

1.32. Since then everything seems to have gone sour for the Conservatives. Mr. Macmillan's carefully planned tactical plan has been completely disrupted already before General de Gaulle vetoed our entry. The negotiations with the Common Market were so long drawn-out by the French tactics at Brussels that Britain could not have joined the Common Market before a general election had to be held—at the latest in the autumn of 1964. Thus the Labour Party was given a possibility to plead for the status quo and attack the Government's policy which implied a change which seemed risky in the extreme. As the traditional conservatism of the British voter is by far his most important quality, this would have put the Conservative Government into a definite disadvantage. It would have been the Government which proposed a risky change and the Labour Party which defended the status quo.

1.33. The French technocrats were extremely skilled. They faced the British with unanimous pressure by the Six on certain vital points, on which the Conservatives were at great disadvantage: British agriculture; imports of temperate zone food from the old Dominions; pledges to the Free Trade Area. On all these points the French were able to mobilise their partners. And the British Government would have been at gravo risk of not being able to carry their own party in the fight for ratification. They were saved from this by de Gaulle.

1.34. The General's brutal interruption of the talks has, to some extent, come to the rescue of the British Government. In this way they could claim that the negotiations had been on the point of success without surrender and prejudice to our pledges when they in fact had been violated, and their violation would have caused extreme trouble in the Tory party.

1.35. In the meantime the Government's policy on the home-front also ran into trouble. Despite the reforms at the Treasury described above, speeches by the Chancellors in the past eight months, and the testimony of Professor Cairncross, their economic adviser, made it clear that the Government's official experts expected a renewal of inflationary tendencies before the end of 1962—hence the brake was kept on economic activity. They were sadly mistaken. Mass unemployment reappeared and with it political agitation for Government intervention. By the time that Professor Paish, Mr. Lloyd's *eminence grise* at the Treasury, was beginning to feel comfortable, with unused productive capacity reaching the 20 percent mark, the pressure to reflate the economy was mounting. Ministers were discovering that their attempt to ensure stability without planning—by deliberately causing unemployment—was politically unacceptable. Hence, even before the end of 1962, fumbling ministerial hands were beginning to relax the brake.

1.36. The retreat from deflation has since then become a rout. Mr. Maudling now suddenly favours expansion, but cannot get it. He has the support of the National Institute for Economic Research for a whopping reduction in indirect taxes—notably purchase tax on cars. This is the easiest and politically most profitable method of stimulating the economy. But its long-term consequences are serious: by throwing more cars on our congested roads, we shall be forced to devote more and more of our limited investment resources to such non-productive purposes as roads and town-planning, and at the same time shackle our economy more and more tightly to the uncertain fortunes of the car industry.

1.37. The Treasury thinking behind this new policy seems to be the reverse of the old assumption that the way to increase exports was to restrict the home market; the new theory is that a rise in home sales will boost exports too. This belief is as foolish as the first. There are only two ways to increase exports—by holding down costs (especially wage-costs) or by devaluation. And devaluation without an effective incomes policy is itself only a temporary remedy. But an income policy, as the unions have made it clear, is impossible without overall planning to ensure a fair distribution of the burdens and an acceleration of economic growth.

1.38. It is at this point that the concept of the National Economic Development Council also recoiled on the Government. The Secretariat had in 1962 presented a statistical picture of the possible development of the country in the next ten years, at that point still assuming that Britain would join the Common Market.

1.39. This picture was quite ambitious and favourable, although perhaps somewhat perplexing. Gross domestic production was assumed to increase by 4 percent per annum. This was of course in accordance with the directive and must not be taken as a forecase and even less as a target. On this basis the Secretariat came to the conclusion that an overall increase by almost £3,000m. per annum over five years in consumption or 3.2 percent per annum could be contemplated with Government expenditure increasing by about £1,000m. per annum overall or by 3.6 percent per annum. This showed partly the immense gain which can be obtained if dynamism were restored to the country, and also disproved the main argument used at the General Election by the Conservative Government about the impossibility of financing social services on a "lavish scale" once growth was accelerated.

1.40. The overall increase in investment was put at £1,700m. per annum over five years, or 6.2 percent per annum. This implied that growth could be accelerated from 2 to 4 percent per annum of the national income by an increase of only from 18 percent to 20 percent in gross investment. Between 1957 and 1961 investment increased from 16 percent to 18 percent. The increase in national income did not accelerate. The incremental capital output ratio was 7.5. It is to be 4.8. This was rather optimistic and could only be explained by the fact that the Secretariat assumed that the substantial capacity unused at the moment would be brought quickly into production. While this represents a welcome break with the Paish doctrine, according to which any increase in the rate of capital utilisation would produce severe inflationary disturbances, it was very difficult to square with the assumption that imports were going to increase only by 4.5 percent, while it was also hypothecated that we join the Common Market.

1.41. The main problem that arises is that it is difficult to believe that, on the basis of complete *laissez-faire*, even the increase in the utilisation of capital could yield the contemplated increase in production. In the past few years a very large portion of total capital investment has been used for less than immediately productive investment and it is difficult to imagine that this would suddenly and substantially change. Certainly there is no indication that there has been a substantial change lately in the composition of capital investment towards productive types and, without discriminating policies, none can be expected.

1.42. The NEDC statistical exercise is obviously of the utmost value as it gives the picture of what accelerated growth, consistent with stability, ought to look like. Its practical credibility is impaired by the fact that NEDC is completely outside the Government machine (the Treasury, for instance, refused to furnish it with an estimate of the increase in Government expenditure) and that, therefore, its blueprint is—at the moment of publication in 1963—totally unrelated to actual Government policy.

1.43. The Government, however, accepted the targets—at least ostensibly—though the statements of its Economic Adviser on policy totally contradicted them and, by insisting on a one-sided wage-stop, made their attainment far more difficult. This contradiction was made even more apparent when the plan of implementing them was published—without authority. It suggests a number of radical reforms, the establishment of a compulsory National Redundancy Fund, reorganisation of training and education towards the attainment of higher productivity, a housing drive to encourage the mobility of workers, and a tax on wealth. As against this it demands a restraint on wages which the unions still categorically reject.

The NEDC plan is weakest, however, on the attainment of balance of international payments. It rejects by implication both severe import restrictions and devaluation, hoping that wage restraint itself would increase the buoyancy of British exports. This aim would be rather audacious even in extremely favourable economic circumstances. But are the circumstances going to be favourable? Bilateral planning of foreign trade, the most obvious solution, is still frowned upon. In any case it is difficult to believe that it will be possible to restrict the rise of imports to the level contemplated by NEDC without some drastic measure of wage restraint. Yet either devaluation or import restriction would raise difficult problems for income policy. There is continued agitation for further cuts in tariffs. These would further aggravate the British international position, as they have done in the past.

1.44. Thus both the target for investment and for the foreign balance obviously call for a very determined effort to keep consumption rising at a rate less than the increase in productivity. This need indeed is foreseen. Productivity is to rise 3.3 percent per annum consumption only 3.2 percent. The achievement of this is inconceivable without a progressive taxation policy so as to obtain most of the increase in investment capacity from the rich, and to obtain close co-operation from the Trade Unions. There is nothing now in British Government policy to suggest that such a complete change in attitude is impending.

1.45. Thus one must come to the conclusion that the present attempt at planning will hardly be successful. It might still serve the Government in discrediting the Trade Unions and thus the Labour Party. But it will not, because it cannot, deal with the problem of restoring economic dynamism, while safeguarding stability. Planning cannot be separated from Government; and it must be done by the Government at the highest level—under the Prime Minister and Cabinet, and it needs experts, such as we do not possess at the moment in the Civil Service. The present Government, the present bureaucracy cannot provide the economic basis of a British renaissance.

2. PLANNING IN A DEMOCRATIC SOCIETY

2.1. The renewed ascendancy in Britain of the notion that a deliberate planning of economic development was required has come surprisingly quickly after the return to the "free market mechanism," itself caused by the failures of the Labour Government of 1945–51 to secure a balanced growth of the economy. Planning was a victim of the unnecessary crises of 1947, 1949 and 1951. Its resurgence was due to the fiasco of the "free system" which brought to the fore, once more, the mounting dissatisfaction with the theoretical basis of the latter.

2.2. This criticism was directed in the first place and with a mounting impatience at the way of life provided by the "free" system despite the increasing capacity to produce goods and services. Apart from the impact of advertising, the increasing concentration of economic power will interfere with effectiveness of the system. The less than full use of resources was felt—unlike in the interwar period—less keenly. It was the inability of the system to combine stability and full employment which came to be at the centre of discontent. Finally—and especially in Britain[4]—there was a growing realisation that the "free" market system does not promote full growth. Neither are sufficient savings forthcoming—indeed not even a level is attained which would be willingly made if individuals had full information about the capacity of the system make use of it; nor is it certain that, if they were, private enterprise would or could make utilise them to the utmost. In all these matters the newly emergent strength of the totally planned economies represented a compelling stimulus to thought.

2.3. Before the war the neo-classical theory of equilibrium ruled supreme. Such dispute as there was among economists about collectivist economic systems centred on the problem of whether such systems *could* at all plan rationally. The defence of the academic socialists (Marxians were not accepted into the fold and their writings were disregarded) against the charge of irrationality of collectivist systems made by Hayek and others[5] was to demonstrate, mainly on the basis of Barone's analysis, that a Socialist Planning Agency could perfectly easily make use of the price mechanism and thus obtain "optimal solutions" based on the free consumer's choice. In this way, through trial and error, economic calculus could be introduced which would make the collectivist system function not unlike a perfectly competitive capitalist economy. The consumer would then be restored to his sovereignty and an optimal allocation of resources would be obtained.[6]

2.4. The underlying implication that a perfectly competitive system would produce optimal results was not doubted by many,[7] though in fact there would be some blemishes reducing its effectiveness. The only question was whether this

---

[4] And in less developed countries with a pressing problem of finding employment to an ever-growing population.

[5] e.g., *Collectivist Economic Planning*, London, 1935. See Schmupeters brilliantly witty rebuttal *Capitalism, Socialism and Democracy* (2nd ed. New York, 1947).

[6] e.g., O. Lange, *On the Economic Theory of Socialism*, Minneapolis, 1938.

[7] This certainly is still the basis of what one might call Social-Democratic or Labour thought on this matter. Cf. Douglas Jay, *Socialism in the New Society*. Mr. C. A. R. Crosland in his writings (e.g. *The Conservative Enemy*, London, 1962) takes a slightly more sophisticated line by trying to discuss the objections against the actual working of a free market economy.

(relative) perfection could be imitated by collective economies. Nor was there any hesitancy about the meaningfulness of concepts such as consumer's sovereignty. The dominant economists thought, after Professor Marshall, that they had an unassailable scientific assumption that the economist is, as an economist, precluded from looking beyond the totality of demand as it manifests itself on the market, from enquiring into the ways in which this demand was formed or evaluating its motivations. The obvious objection that in this way they would give their non-scientific, non-objective, political blessing to the *status-quo*, did not seem to have occurred to anybody.

2.5. The reaction to this attitude was acute after the war, and it came from two directions. On the one hand, the foundations of welfare economics were undermined, by showing that they inevitably involved value judgments, i.e. political assumptions. On the other hand, and culminating in Professor Galbraith's book *The Affluent Society*, the assumption that wants were given was questioned. They could be created and modified. This conclusion had momentous consequences. If some systems create more wants than others, how is their relative merit to be judged? Even if one was more "productive" than the other, in the sense of producing a greater real output valued at some system of prices, it might involve a net dissatisfaction, disutility, inasmuch as it might create more wants than it was able to satisfy relative to a less "dynamic" system. Traditional economics could then only be saved by assuming that the creation of wants was itself to be the way to *superior* living, that it represented a good in itself, progress through the creation of divine discontent, because it represented an enlargement of the field of choice which had to be accepted as a boon without any further discussion and *without this assumption involving a political or value judgment.* It is the new "activities" which give rise to new wants. The whole process is in a sense spontaneous, ingrained in economic change, which it is not the economist's job to question. After all, it would be a presumption to claim that the man from Whitehall or for that matter Harvard University, would know better than other people what is good for them.[8] This is an intolerable arrogance, and Veblen, Galbraith and other critics of high-pressure advertising were roundly condemned for it.

2.6. This effort of rationalising with "hurray"-words, the arbitrary application of a system of thought appropriate to the problem of the optimal satisfaction of *given wants* based on *unchanging preference* to the basically different problem arising out of changes, and especially *biased changes in wants* has not been altogether successful. Once the analysis of social unbalance, of lagging public expenditure and the nagging doubt whether the free market system can cater for social priorities has been raised, it is difficult to suppress. With it goes the definition of progress on the basis of production alone, the functional test of the performance of an economic system, the output standard, the valuation of real income on the basis of the output at some conventional price. The criterion of economic progress must then be redefined in political terms on the basis of social priorities which consciously take into account values such as security of the job, absence of uncertainty, lessening the arduousness of work, increased leisure, for which the price system does not cater.[9]

---

[8] It is significant that it was the right-wing supporters of the Labour Party who attacked a recent Report (the so-called Pilkington Report) which expressed doubts about the effect of commercial television advertising on public taste and mental health.

2.7. It is probable that the attack on the meta-economic basis of neo-classical equilibrium theory would not have come if, to the annoyance and confusion of its protagonists, the productive performance of the Soviet economy, valued in the conventional way, had not proven superior to that of the economies with decentralised decision-making. Comparisons made on the basis of the "new" welfare economics have turned out to be to the credit of the planned economies. It was this, and no sudden internal enlightenment, which forced economists outside the Russian orbit to re-examine their approach to the subject. It was the Russian productive performance which put growth and development back into the centre of interest of economists, from which it had vanished for 150 years.

2.8. The reaction to this by the less thoughtful anti-communists was simple. They merely imitated the Stalinite antics in showing that the conventional methods of economic calculus could give the required anti-Soviet answers by either discounting the rapidity of Soviet growth[10] or showing that it did not fulfil any purpose. "Conspicuous production" was the explanation of the Soviet success, while conspicuous consumption, or the creation of new wants in the capitalist orbit which menaced conventional economics, was disregarded or minimised. In actual fact the Soviet system, by being able to regulate the production of consumer goods and the rhythm in which new wants were admitted and catered for, *could* more easily ensure a balanced development than could an individualist system based on the profit motive. Whether it *will* in fact do so is one of the great questions of the day. The tendency of the high command of the Communist to set their targets in terms of American achievements does not disclose either much sociological insight or socialist thought. It may well be that the Soviet will develop towards an imitation of the worst features of subtopia combined with far greater compulsions to conformity than the United States have to suffer. All that can be said is that this deformity certainly is not implicit in the economic mechanism of the system—though it might be in the sociological framework of the dictatorship.

2.9. American efforts to discredit the purely economic achievement disclose more of their own weaknesses rather than those of the Soviet. In any case the silly trick to demonstrate Soviet inferiority by showing the time-lag behind the U.S. has already been overtaken by events. So long as the Soviet production was below the American level reached in the period 1915–29, when expansion was relatively fast in the war and the New Era, this method yielded impressive results. Unfortunately it meant that the Soviet now appears to be overhauling 10-15 years of American progress each year, for between 1929 and 1939 the U.S. hardly grew at all.

---

[9] This is especially important in countries (as e.g. in India) where a (large) section of the community is still living at the margin, where the creation of wants among the well-to-do minority necessarily impinges on savings and thus on the capacity of the economy to provide employment in the "modern" sector. The insane idea of launching a "people's car" (of an output capacity of some 30000 p.a. with the "people" numbering some 440 odd millions) is a typical instance of the total lack of realisation of what is involved.

[10] There was of course a legitimate task (especially in the Stalin era) of correcting Russian production indices which overstated the speed of advance.

2.10. A response to the threat of the Soviet challenge cannot be made by such legerdemains. It could only be based on a detached analysis of the comparative advantages of the two systems and an intelligently co-ordinated effort to strengthen the advantages as well as to eradicate the handicaps of the economies based on decentralised decision-making.

2.11. This analysis has to proceed at three levels.

(a) Of these the first concerns *the allocation of existing resources*, i.e. the efficiency of the system as it is. This was the problem which was almost exclusively at the centre of the pre-war dispute, and was thought to be resolved entirely in favour of the market system.

It is of course not easy to make much sense of the meaning of the notion of reallocation of *given* resources of which the most important of the non-human ones, capital, is to a very large extent specific and non-transferrable. Thus it is in itself difficult to conceive of *reallocation*, conceptually, independent from technical progress and growth. Even if it were, it is clear that it cannot, in the longer run, (when we are not all dead yet) exert anything like as important an influence as factors which are *cumulative and compound* in their effect. So far as these are concerned there is

(b) the handicap imposed on a completely non-planned system by the *functional defects* of the system by the less than full utilisation of resources. A *fuller use of resources*, faster innovation and higher investment, i.e. factors which influence the rate of growth of the economy, are then seen as the decisive ones in assuring the success of a system in the conventional sense of the word.

(c) Finally there is what one might call the *long-run structural defect* of economies without a coherent plan. This stems form the inability of the non-planned system to devote a *sufficiently large proportion of their total output to investment*, a proportion which might well be allocated voluntarily if both savers, and those responsible for investment were fully aware of the implications of the decisions on their own fate and that of the system.

2.12. The first, the allocative function of the system, its microeconomic effect, is perhaps the least important. When all is said, remedial action on this plane would only result in a once-for-all gain while the other two causes represent a cumulative threat. This estimate is confirmed by the superior economic growth which has been achieved by the Soviet, measured in conventional terms. There can be little doubt that the Soviet economic calculus, the choice of techniques and between projects, was gravely deficient in the past. These shortcomings of the Soviet planned economy were due partly to the neglect of the true opportunity cost of various factors of production, especially capital, and partly to their inability to achieve a balanced growth as between agriculture and industry. It is the crushing superiority on the front of industrial dynamism, by being able to achieve a more or less full utilisation of resources and to devote a calculatedly higher proportion to investment and innovation which was responsible for their progress. Its impact far outweighed the retarding effect of the inappropriate distribution of resources among productive tasks.

2.13. The question then is whether the non-totalitarian countries can achieve some compensation by overcoming their basic difficulties. This is necessary because several important structural changes in the non-Soviet economies have contributed to the *weakening of the advantages* inherent in *self-reliant decentralised management* based on the profit motive in their day-to-day decisions. Thus if the non-Soviet world is to meet the Soviet challenge, a strengthening of its growth potential is in order.

2.14. The first of these factors weakening the efficiency of the price mechanism is the impact on the functioning of the signalling system, essential in a decentralised economy for the assurance of a co-ordination of decision, as a result of the *concentration of economic power.* The second is the problem of *risk* and its influence on investment, and the third is the problem of the *increasing importance of indirect gains and losses* which ought to, but cannot be taken into account by private entrepreneurs in their decisions, gains and losses apart from the impact of the fuller utilisation of resources and the problem of accumulation.

2.15. The weakening of the effectiveness of the signalling system has first of all arisen because of the fact that costs are not independent of demand, and the entrepreneur in many cases will not know what his precise costs are. Cost is not independent of output, and output is not independent of sales policy. Thus we are confronted here with an indeterminacy worse than would be experienced by a totally planned system where selling costs can be regulated from the viewpoint of the community as a whole.[11] Moreover, working for known clients, it would not be in the firm's interest to try to snatch maximum profits in the present if this will prejudice their position in the long run. Thus determinate pricing on the basis of a perfect market goes. No doubt demand can be gauged pretty accurately in the short run. It is also probable that it will be fairly unresponsive immediately to changes in price. But the firm has to stay in business. It must, therefore, take into account the consequences of its own policy, as well as that of its competitors, on its relations with its clients. Competition among a few firms is more like a game of bluff than a mathematically determined play. All this makes decisions subject to discretion hence to judgement, and risk of mistakes. It certainly can no longer be claimed to be unequivocally determined by the inter-play of impersonal market forces.

2.16. There is a further difficulty. In the longer run adjustments will take the form of new or modified equipment. Thus changes here will not be reversible. Once the new equipment has been installed there will be no return to the original situation. Some equipment will be less productive but more flexible. There is then a choice between various techniques which depends on the judgement of the future, and especially on the stability of demand for a specific product. Thus the simplicity of a determinate signalling system, which will result in optimal decisions is destroyed in this respect too.

2.17. Knowledge could in a static theory be taken as granted. If nothing changes and everything recurs, knowledge is perfect. Uncertainty and risk will prevent decisions and thus retard adjustment to new situations. Secondly, in modern conditions firms will certainly have regard to the actions of their fellow competitors. Their prices and production are not fixed in the sort of world markets—like agricultural products—which any one producer cannot influence.

---

[11] *i.e.*, either uniform treatment can be adopted, or discrimination on the basis of some intelligible principle.

2.18. It would in fact be a perfectly competitive behaviour that would be irrational in this system. The single firm will not raise its price because it will fear to lose its markets; nor will it initiate the price war by cutting its price. The contrast in the changes in the U.S. steel prices in 1954–5 and 1962 respectively, the unanimity of action by all concerned despite anti-trust legislation, the uncertainty of touch displayed, and the lack of consideration of wider issues on the part of the industry on the one hand, yet the conclusive effect of their decisions on the general movement in prices, all this exemplified the importance of the new forces which have put an end to a determinate price system which expresses fully the socially relevant relative opportunity costs.

2.19. If, however, firms have to look over their shoulder to see what their competitors are doing, if they are unwilling to slash prices to run their plant at full capacity, and if, moreover, plant capacity is determined largely by technical considerations, there will be a substantial amount of excess capacity. Part of this capacity will in any case be required in an individual private enterprise economy by *each* firm for insurance against not meeting possible specific orders and thus prejudicing the individual future. An individualist economy will therefore, in principle at least, have to carry a great deal more excess capacity than a centrally planned one.[12] In addition, and especially in manufacturing, prices will not be fixed uniformly on the basis of cost, thus further distorting the accuracy of the signalling system.

2.20. The growth of firms has another important consequence. The picture of new entrepreneurs forcing their way towards the top, innovating, opening up new ways, is a very attractive one. It has become outdated with size.[13] Marshall, following Adam Smith, stressed this imponderable aspect of the problem, and rightly so. In large firms, as in a state bureaucracy, innovators are not at a premium. They are uncomfortable to live with, not what the Americans call organisation men. The character of the oligarchy ruling oligopoly is deadening rather than invigorating, and it is only exceptionally that increased activity can be forced upon them by the outsider. In this respect the great basic advantages of the system with decentralised decision making are relatively on the decline.

2.21. In all these matters, therefore, *a co-ordination of investment decisions*, a reduction of the risk involved in being prematurely expansionist, might pay handsome dividends. *Planning, in the sense of indicative planning, thus is justified on the score of the weakening of the automatic co-ordinating system of the price mechanism.*

---

[12] This can be—and has been—offset, and more than offset, by failure in planning and bottlenecks in supplies. The hoarding of components was notorious and still remains one of the main handicaps to efficiency in the Soviet system.

[13] In point of fact, the greatest innovators in Europe have been the nationalised Volkswagen and the nationalised E.N.I., both managed by geniuses unlikely to have come to the fore in a different framework.

2.22. Far more important, however, are the dynamic and cumulative shortcomings of the system also tending to strengthen this conclusion.

2.23. *A full use of resources* in a non-planned economy activated by the profit motive creates an atmosphere in which wage bargains become a customary round. The lack of co-ordination between these bargains, the need of trade union leaders (in order to keep the loyalty of their followers) in equalling the performance and, if possible, scoring off their colleagues, all this imparts an inflationary bias to a decentralised economic system in a framework of modern industrial institution. Attempts have been made on the basis of completely unscientific assumptions and methods to quantify this relationship in order to prove that slight unemployment would be sufficient to restore balance, to recreate the pristine simplicity of monetary equilibrium.[14] These attempts have failed miserably. It has been shown[15] that the militancy of the trade unions is by far the most important factor in determining the rate of increase in money wages within a wide range of monetary conditions. Single firms or even single industries could not resist the bargaining tactics of labour which in days, when full employment had not yet been forced on governments, could easily have been crushed.

2.24. No doubt it would be possible, to ensure monetary balance by deliberately creating large-scale unemployment. But the consequences of such a policy on the progress of the economic system would be fatal. For if uncertainty is deliberately created about the future of demand—and this would be necessary in order to encourage individual firms to resist the demands of trade unions—it would also involve a slowing down of growth. In the end it would threaten even technical progress: trade unions would resist innovation on the plea that it created unemployment. Only a full use of resources, leading to faster innovation and higher investment, can secure to the decentralized economies a rate of expansion comparable with that of the centrally planned system. It is precisely the dissatisfaction with the alteration of deflationary and reflationary measures, that has been responsible even among Conservative adherents of the "free" market system for the support for long-term planning.

2.25. It is of course true that in economic systems which, for one reason or another, were able to weaken or eliminate trade union organisation, stability could be assured, yet in the longer run very substantial acceleration of the growth of real wages could also be secured. Ex-post, trade unions would have been justified in accepting a redistribution of income towards profits because by higher investment this produced a faster increase in real wages than would have been attainable if greater Union strength would have enforced premature increases in money wages. The large agricultural reserve army in Japan, mass-unemployment in Southern Italy, the immense immigration from Eastern Germany to the West, the failure of the general strike against de Gaulle, provide interesting examples in which this in fact happened.

---

[14] e.g, the so-called "Phillips" curve. Cf. K. Knowles and C.B. Winsten in the *Bulletin of the Oxford University Institute of Statistics*, 1959.

[15] Sir D. MacDougall, *Economic Record*, 1961.

2.26. It is not a stable situation. Once full employment works through the system the atmosphere changes and trade union activity becomes prominent, unless it can be suppressed by forcible police action, i.e. the end of democracy. In a democratic system, however, stability and full employment could only be combined by demonstrating to trade unions in advance the great benefit of their compliance with a policy of combining wage bargains with some regard to increase in productivity, while at the same time assuring some justice in the distribution of income, and not permitting profits to benefit without restraint from the concession of the wage earner. This would be possible if a large part of the increase in the savings needed were to be arranged by collective saving, i.e. budget surpluses from an equitably distributed taxation system. This would not necessarily be unfavourable to the capitalist, for the more rapid increase in national income would result in a vast increase in total profits, even if their *share* were not as high as it would work out without a plan. Here again a co-ordination over the long period is needed in order to restore the effectiveness of the price system to its theoretical level.

2.27. There is, then, the somewhat related but long-range problem of growth, the problem of the determinants of savings, investment and innovation in a given society. In one sentence, the planned economy in this respect seems to have a twofold advantage, because on the one hand it can take into account in its calculations the indirect gains resulting from the new level of investment decisions, and on the other it can deliberately make available to all the advances in techniques which individual entrepreneurs must keep to themselves if they are to be able to pay for the research they have to undertake to gain the new knowledge. *A planned economy thus can externalise internal economies and internalise external ones.*

2.28. In addition to the extra gains which can be made by single industries, or by a restricted number of industries linked together, from the increased scale of their development through increased productivity and returns, there is thus a more general source of what might be called general external economies due to an accelerating growth even in fully industrialised countries, and much more so in underdeveloped ones. I hope to treat this vital problem in detail in a different context. Here I shall confine myself to a summary of the argument.

2.29. Neither individuals nor even corporations can fully realise the potential gains which would occur if their investment and growth were speeded up, because their successful realisation would depend on other people's action, especially on their investment and savings decisions. Thus there is likely to be an objective failure: the rate of growth might fall to a lower level than it would be if all concerned knew the reciprocal consequences of each other's action and could plan them jointly.

2.30. On the investment side entrepreneurs settle down to a decision-making pattern which is influenced, if not determined, by the historical level of their previous investment which proved successful. This follows from the fact that those controlling the majority of investment will decide whether all other investment (even if well-planned) will be successful. One investment justifies another. The 'marginal productivity schedule' is not fixed. Investment does not "exhaust" opportunities except in a short-run sense. It does not "shift" upwards automatically either. The historical upward shift is the result of coincident investment decisions. If a majority refused to invest, those who would, would lose out. The penalty of over-eagerness is failure and losses.[16–17]

2.31. If investment plans are not coordinated no rise in investment will be expected and none will then take place. In order to create the conditions of high domestic profit and progress, the opportunities have first to be created. Paradoxically, the existence of an overall plan might be needed, precisely in an economy not subject to central planning, to speed growth. Otherwise risk might increase and opportunity be wasted.

2.32. Equally important is the incapacity of a system based on decentralized decision-making to match increased investment by increased savings, in the sense that no cumulative inflation originates[18] and the attempt is frustrated by a premature increase in consumption. Ex-post (as the Japanese and German case clearly shows) Trade Unions realise the benefits of increased investment and growth. But ex-ante they cannot connive at what would in fact have to be a redistribution of income towards profits, especially if the policy has to be inaugurated (as in Eisenhower America and in Tory Britain) *after* decontrol and the free-for-all has been made the corner-stone of Government policy—in the name of accelerated growth! Only a national compact, backed by appropriate tax reform might be able to perform this task.[19] But, until Mr. Kennedy's victory, the trend was contrary to such solutions; and even his Government has in this respect been paralysed by Congressional opposition. The German, Japanese, French and Italian cases demonstrate *a contrario*, each in its way, that under a system of decentralized decision-making the inauguration[20] of a regime of high investment and growth depends on some special factor paralysing or weakening Trade Union bargaining power and the possibility of achieving expert surpluses without inviting retaliation.

2.33. Much has been made, especially in the immediate post-war period, of the difference between financial and physical planning. Physical planning was extolled as something especially praiseworthy, while reliance on financial incentives and penalties seems to have been equated with the profit motive, and disparaged as rather more vulgar.

---

[16] The consistent unscientific propaganda of the I.M.F. suggesting that a 3 percent growth rate —grossly inferior to the Soviet one—is "realistic", might have decisively contributed to the slowing down of growth in the non-Soviet orbit. It has done the work of Mr. Khruschev.

[17] This view is in no way inconsistent with the fear that in *weak* countries the expansion of investment in the rich countries is likely to cause havoc (except in the case of endowment with natural resources the demand for which rises with increased investment). The problem is one of the likelihood of complementary rather than competitive (and destructive) investment. It is my contention that these probabilities are not symmetrical in weak and strong countries.

[18] Mr. Kaldor's recent series of models of growth (e.g., *Economic Journal*, December 1958) seem to assume that Trade Unions do not exist and that increased investment thus would lead to increased saving and not to cumulative inflation in the end destroying itself.

[19] Cf., Productivity and Inflation, *O.E.P.*, June 1958. The majority of the experts appointed by O.E.E.C. have agreed on the need for a wages policy. They have, however, not discussed in detail its political preconditions. Cf. *The problem of rising prices.* W. Fellner, Milton Gilbert, Bent Hansen, Richard Kahn, Friedrich Lutz, P. de Wolff, O.E.E.C., Paris, 1961.

[20] Not necessarily the continuance. Cf. Productivity, *loc. cit.* There seems to be no close correlation between money wage demands by Trade Unions and the rate of growth achieved. Thus a high rate of growth might enable the devoting of an increasing part of national income to investment, though we have no successful examples of such harmony without breaking the Trade Unions.

This obviously is quite unjustified. Planning in physical terms is essential in preparing an over-all co-ordination of investment projects in order to provide the basis of the modernisation of industry and the maintenance of an even employment situation while safeguarding growth. Once the physical target figures for the basic industries have been arrived at it depends entirely on tactical requirements whether or not the implementation is to be done through physical controls, financial incentives, or (preferably) a mixture of the two.

2.34. The failure of the first attempt in Britain—after 1945—to arrive at a satisfactory system of planning was that the relationship of physical to financial controls was never understood, and that physical controls had, therefore, to be used to counteract the market forces. Yet subsidies and tax measures could have modified the market forces to help in achieving the aims set by physical planning. It is obvious that so long as public ownership is not all-embracing, physical controls acting contrary to the profit motive will always be subject to the risk of evasion. Thus suitable modification of market forces through fiscal measures permits a much more flexible and less administratively arduous implementation of basic plans.

2.35. Admittedly in some cases the gentle pull and push of these forces will prove insufficient. This is especially the case if full employment is maintained and, therefore, the normal safe-guards through price stability, or relative stability, are suspended. The extent to which direct controls will have to join fiscal measures will depend on the width of public ownership and the degree of concentration of the private sector, that is, on the effectiveness with which agreed industrial investment plans can be expected to be implemented automatically.

2.36. There are two main domestic fields—apart from foreign payments —in which the success of fiscal measures must be expected to be relatively low, both of them of increasing importance for a balanced economic policy in advanced nations and in those underdeveloped ones which hope to achieve a rather rapid growth.

2.37. The first of these is the location of industry. This has been increasingly dominated not so much by consideration of costs and profits (except perhaps for heavy industry) but by consideration of convenience and psychic income for management in choosing congenial places to live. It has proved difficult to modify location decisions by fiscal means in most countries of the world. The relative failure of the Italian Government in the south is one of the most impressive proofs in this respect, for lavish financial provisions, great knowledge, immense enthusiasm and idealism, and intense propaganda were deployed. Licensing of investment is accordingly required.

2.38. Another domestic field hardly less important now, and still increasing in impact, is the field the incitement for higher consumption to maintain profits. This is achieved by the introduction of new products, and especially by the continuous modification of products to push physical obsolescence. In this direction, too, fiscal encouragement or discouragement cannot be expected to do much. A control of advertisements, a strongly organised campaign of consumer enlightenment, a toughening, through publicly owned communication channels such as television, of the public against advertisements emphasising the fundamentally unsound aspects while leaving the spread of information unchallenged, would be the only way in which a less competitive framework could be made compatible with private enterprise. In the end only socialism, social ownership, can deal with it.

2.39. These controls could be backed up by discriminating financial incentives. The process of technical renewal could be encouraged by providing additional investment allowances on a moving scale for increased efficiency, while penalising the continued retention of old capital by disallowing depreciation allowance if the renewal does not take place within a certain prescribed period. Control of location could be assisted by a discriminatory handling of depreciation and investment allowances, both as to location and as far as branches of the economy are concerned. If, for instance, luxury commercial building is to be discouraged, negative investment allowances could be introduced, i.e. whole of the cost would not rank for depreciation. In this way the work of the direct control will be helped.

2.40. Apart and beyond these the most essential of the direct controls are those on foreign payments and trade which eliminate the threat of a conveniently originating financial crisis imposing a veto over policy and speculative attacks being made against the currency, which, being cumulative in character, might create the instability it fears.

2.41. The fully developed democracies seem to face somewhat paradoxical problem. On the one hand both the objective world position and the drift in their economic development seem, as we have tried to demonstrate, to demand a more conscious policy and active initiative on the part of their Governments in the determination of the growth in national productive capacity and the allocation of its fruit to avoid instability and a relative decline in the strength of the non-Communist world.

2.42. Yet these Governments seem ill-fitted for these tasks; and the more conscious anyone is of the need for planning and conscious direction of our economic destinies, the more inevitably conscious he becomes of the shortcomings of the organs available to carry it out.

2.43. In a perceptive analysis the late Mr. Bevan recognised, as early as 1950, the deep-seated character of this problem, even though he then regarded it as restricted exclusively to the underdeveloped areas of the world.[21]

> "It is highly doubtful whether the achievements of the Industrial Revolution would have been permitted if the franchise had been universal. It is very doubtful because a great deal of the capital aggregation which we are at present enjoying are the results of the wages that our fathers went without.
>
> It is not easy in some parts of the world where life is exceedingly primitive to entrust to all the people the power of democractic selection, and then to ask them to forego some immediate delights. It is not easy."

2.44. He later advanced far beyond this narrow interpretation. In a speech which excelled in its oratorical brilliance as much as its analytical content he said:[22]—

---

[21] Democratic Values. Fabian Tract No. 282, London, 1950.

[22] Hansard No. 486.

6

> "I would describe the central problem falling upon representative governments in the Western world as how to persuade the people to forego immediate satisfactions in order to build up the economic resources of the country. Let me put it another way. How can we persuade the ordinary men and women that it is worthwhile making sacrifices in their immediate standards or foregoing substantial rising standards to extend fixed capital equipment throughout the country? This is the problem and it has not been solved yet."

3. CONCLUSIONS

3.1. It is thus clear that planning is required even in a democratic setting, if only to ensure stability in the system when resources are fully employed. The tendency towards decontrol has led, and is bound to lead, either to monetary instability, or to a far less than full employment of resources, and consequently to a slower rate of growth than would even be implicit in the voluntary decisions of the individuals making up the community had full employment been secured. There is thus here a fatal *functional defect* in the system. The elimination of this defect requires conscious co-ordination of investment decisions, of income distribution, and the elimination of disturbing cumulative speculative excesses.

3.2. If this is probably the most vitally urgent case for planning, it is obvious that the planning implied by this requirement alone does not really go very far. It need not fundamentally interfere with the essence of the individualist system based on private property. No doubt income distribution would have to be made more equal than it has been in the recent past to secure the co-operation of Trade Unions. An income policy implying some wage compact coupled with some restraint on profit margins through powers to control prices (which need not really be exercised but must be available), would probably do the trick. If full employment were to be assured in the non-Soviet orbit as a whole, and if entrepreneurs were to be convinced that stability and full growth can be combined the elimination of risk of unexpected losses due to deflationary periods would probably offset (and might more than offset), the inimical impact on incentive to risk-bearing and enterprise of a less unequal distribution of income.

3.3. Should the elimination of these risks, moreover, result in a substantial acceleration of economic growth and incomes, this by itself would permit the lessening of tension between the various classes of the community because it could provide a satisfactory advance in the standard of life all round.

3.4. To ensure international balance there would have to be some monetary collaboration and some possibility for the conscious creation of international liquid resources eventually through the organisation of an international central bank with powers to rediscount or open market operations, the countervalue of which would rank as an international reserve acceptable to all.[23]

---

[23] I have dealt with this problem in my book on *The Dollar Crisis* (1949) and in an essay published in the Economic Journal (1960) and reprinted in *Unequal Partners* (Blackwell, Oxford) and Einaudi, Torino (1963).

3.5. There is, however, a strong case even in fully developed countries, and an overwhelming case in the less developed areas now in the process of industrial and agricultural revolution, for at least a partial liberation of demand from the unrestricted influence of the profit motive. It is hardly compatible with balanced progress if entrepreneurs are permitted deliberately to create discontent and new wants to secure outlets for the produce. In highly developed areas this finds its expression in the creation of consumer restlessness on the one hand, while on the other an unnecessary degree of poverty is permitted to exist and collective consumption is neglected. The social unbalance so brilliantly analysed by Professor Galbraith is an inevitable outcome of production for profit coupled with a resistance against taxation, especially taxation impinging either directly or indirectly on the upper income strata.

3.6. In less developed areas its impact is, of course, far more devastating. Here the conspicuous imitation of upper class modes of living in highly developed areas, added to the traditional expression of social superiority in terms of abundant services, results in a debilitating decrease in investment and thus in the capacity to provide employment at tolerable rates of income. Coupled with the explosive population problem, this creation of needs undermines stability in the non-Soviet orbit.

3.7. Planning in this sense of liberating the energies of the people towards leisure and its more enjoyable utilisation,the provision of a pleasant atmosphere and environment, a deliberate reduction of competitive tensions, all these are difficult to conceive without a very important extention of some type of social ownership.

3.8. The richer and more successful the system, the greater seems its psychological malaise. It is significant from this viewpoint that dissatisfaction with the social mechanism has been more acute in the United States than in Britain.

3.9. It has been suggested that an escape route can be found from this malaise by a comprehensive system of redistributive taxation. This view has been discredited by the experiences both in highly developed and in underdeveloped areas. In those cases where the owning classes could not find a convenient escape they succeeded in discrediting the taxation system and preventing its reform. Moreover, especially in highly developed areas, an escape was found by increasing managerial incomes so as to offset and more than offset the influence of taxation. The net effect is one of growing inequality.

3.10. In less developed areas the problem could be handled by direct controls rather than social ownership. There the prohibition of imports of conspicuously wasteful durables, and the strict use of licensing laws to prevent the premature production in small quantities of such goods for the delectation of the few, could in fact enforce on society a consumption pattern more in line with the requirements of an accelerated growth of the investment capacity and thus the capacity to provide employment at tolerable income levels. Without any doubt, one of the most disheartening developments of the postwar period in Latin America has been the premature establishment of rather inefficient luxury industries whose very being saps investment capacity and thus stabilises the present dreadful discrepancy in the standard of life of the population. The process is about to be repeated in Africa and in Asia. It has as yet only been checked in India. Even there the so-called "popular car" project shows that people are completely unaware of the task and purpose of public ownership in accelerating and concentrating the nation's economic effort.

3.11. In highly developed areas this solution seems impossible. Here we already have a great array of industries whose products serve to impress one's follow beings. Direct controls could not attain the end of providing a greater social balance. Both equality of income distribution and the lessening of the undesirable aspects of product differentiation could not be attained without direct social ownership.

3.12. There remains finally the problem of ensuring a higher degree of dynamism in the non-Soviet orbit than would be attainable on the basis of the spontaneous unguided decisions of individuals. This does not necessarily mean that forced savings have to be accumulated. The process of planning and co-ordination of investment projects itself, by assuring the immense indirect gains which can be attained, might induce people voluntarily to increase their purposive rate of accumulation. Nevertheless it is probable, if only to attain greater equality in the distribution of wealth and therefore of future income, that a considerable part of the total additional effort would have to be in the form of *collective saving*, and it is desirable that it should be at least partly *additional saving.*

3.13. To some extent this is implicit in the considerations affecting stability at full employment. An income policy, to be successful, must include a policy on profits, and (unless social ownership is expanded far more than now seems likely) this can only be accomplished if a large part of public investment (which in any case would have to increase to assure better social balance) would be financed out of current revenue, mainly falling on the well-to-do. The alternative, as in Germany and Japan, is a violent distribution of income in favour of profits, a process which is not compatible with full employment in a framework of well-organised labour relations. Only if the trade unions are weakened by some economic or political force would this alternative solution be attainable.

3.14. We have thus come to a picture of a partially planned economy in which social balance and a more equal distribution of income is sustained by a considerable public sector. In this sector managerial remuneration can be kept at a desirable level. Conditions of work, the provision of leisure and the avoidance of continuous stimulation of new wants can also be checked. For the rest some controls on the location of industry, and on profit margins will have to be combined with redistributive taxation which will have to be kept, and can be kept, at a high level because under these conditions such taxation will not affect the mainspring of investment decisions. Finally overall foreign exchange controls or, alternatively, international co-operation would keep an external balance without periodically having to depress investment and demand.

I have no doubt that only such a system could attain superiority over the totally planned economies impelled forward by state fiat.

PARAMETRIC SOLUTION AND PROGRAMMING OF THE HICKSIAN MODEL

By RAGNAR FRISCH

Assisted by ASHOK K. PARIKH[1]

Institute of Economics, University of Oslo

THE HICKSIAN MODEL

In a nutshell the Hicksian model[2] can be described as follows, if time is denoted by t (say calender year):

Notation for five variables. $C_t =$ private consumption; $I_t =$ net *induced* investment; $G_t =$ Government use of goods and services on current account; $H_t =$ net *autonomous* investment[3]; $Y_t =$ net national income[4].

Notation for four parameters: $\alpha, \beta, \kappa, \lambda$. Of these α and β are in the main structural parameters not subject to Government decision, while κ and λ may be thought of as being subject to Government decision at least to some extent. The parameters κ and λ will define the Government *strategy* with respect to the total Government expenditure on current account.

Four equations:

Consumer behaviour, $\quad C_t = \alpha Y_{t-1} \quad \ldots (1)$

Behaviour of induced investors, $\quad I_t = \beta(Y_{t-1} - Y_{t-2}) \quad \ldots (2)$

Government strategy with respect to current account expenditure, $\quad G_t = \kappa Y_{t-1} - \lambda(Y_{t-1} - Y_{t-2}) \quad \ldots (3)$

Definition of net national income $\quad Y_t = C_t + I_t + G_t + H_t. \quad \ldots (4)$

Degrees of freedom. As it stands the model has five variables and four equations, hence one degree of freedom. This degree of freedom may be thought of as being generated by the variable H_t, i.e. autonomous investment. This variable may be looked upon either as *deterministically* given (i.e. each value H_t is given for reasons outside of the model) or as *stochastically* given (i.e. the variables $H_1, H_2 \ldots$

---

[1] The Indian economist Mr. Ashok K. Parikh, M.Sc. (Econ.) (London School of Economics) stayed some months in the Oslo University Institute of Economics in the summer of 1962.

[2] J. R. Hicks (1950): *A Contribution to the Theory of the Trade Cycle*, Oxford. See also 7.3 in R. G. D. Allen (1957): *Mathematical Economics*, London.

[3] Hicks uses the symbol A_t for autonomous investment. I prefer to use H_t because it corresponds better to the Oslo notation.

[4] In standard national account terminology the Y_t would have to be called net national *product*. The designation "income" is the national account terminology connected with the concept of factor cost. I have on several occasions strongly criticized the concept as being neither logical nor fruitful. I am therefore quite happy to use the term "income" for Y_t, even though it may have nothing to do with factor cost.

have a given simultaneous probability distribution; possibly as a special case all the $H_1, H_2 \ldots$ may be stochastically independent). In any case the values $H_1, H_2 \ldots$ will have to be considered as *data* when we approach the solution of the model.[5]

This means that the solution should be sought in the form of four time-functions C_t, I_t G_t Y_t explicitly expressed in terms of the four parameters $\alpha, \beta, \kappa, \lambda$, the initial conditions and the values $H_1, H_2 \ldots$.

Method of characteristic roots and method of recurrence. The solution may either be obtained by the *recurrence* method or by the method of *characteristic roots.*

In the recurrence method we start from a given initial situation of the variables and use the model (1)-(4) to compute the constellation of the system in next point of time and so on.

In the method of characteristic roots we start first by obtaining the solution of the associated *homogeneous* systems, i.e. the system that is identical with (1)-(4) except for the fact that the term H_t is *dropped.* In other words one puts $H_t = 0$ for all t. And this homogeneous system is solved in a particular way, namely, by computing the so-called characteristic roots. And by the help of the solution of the homogeneous system one finally derives the solution of the complete system (1)-(4).

It is very useful to have the solutions available in the both forms (cf. equations (35), (127), (130) and (133)).

The recurrence method has an advantage compared to that of the method of characteristic roots because the recurrence procedure yields immediately the forms which the functions in question must *necessarily* have in order that they form a solution. In the method of characteristic roots, we do not get immediately anything more than forms which are *sufficient* in order that the functions in question shall satisfy the system of difference equation. It follows, however, from classical mathematical theories (proved in a rather elaborate way) that under certain conditions the function forms obtained by means of the method of characteristic roots, are actually the forms which the functions must necessarily have in order that they shall satisfy (1)-(4).

In the sequel, I will use both these methods and compare the results.

The system (1)-(4) is a simple one because it can be *partitioned* in such a way that the whole problem is reduced to a consideration of a simple difference equation in one variable namely Y_t. Once the explicit expression for Y_t is obtained, the remaining follows from (1)-(3).

The equation in Y_t is derived by inserting (1)-(3) into (4). This gives

$$Y_t = \mu Y_{t-1} + \nu Y_{t-2} + H_t \qquad \ldots \quad (5)$$

(one equation in two time functions Y and H)

where

$$\mu = \alpha + \beta + \kappa - \lambda \qquad \ldots \quad (6)$$

$$\nu = \lambda - \beta. \qquad \ldots \quad (7)$$

---

[5] There is no particular reason to assume that H_t must be non-negative; it may in principle be positive, negative or zero, but of course in practice it will as a rule be positive.

A realistically weak point in the model. The equation (5) brings to light a realistically weak point in the model. The equation shows that whatever the past course of the economy has been, the net national income Y_t can in any given point of time t be rendered *arbitrary great* if H_t is chosen sufficiently great.

It is, of course, true that frequently the constellation of the economy, and its structure, may be such that an expansion will occur if *hidden resources* are released, either through determinate government action which will stimulate autonomous investment H_t, or through more or less accidental happenings (inventions, wars etc.) which will also stimulate autonomous investment. But I think that Western economic thinking on trade cycle regulation and on economic growth is all too one-sidedly concentrated on this point. The present model where H_t can *ad libitum* produce an expansion in Y_t is an example. H_t is by (5) so to speak an extra addition which we may, if we like, give to Y_t when the point of time t has been reached.

The analysis of the role of autonomous investment (or more precisely of autonomous investment *starting*) ought to proceed in a more thoroughgoing way by relating the possibility of autonomous investment to the existing production capacities and other *bounds* in the economy. In order that the additional term H_t in (4) and (5) shall have a meaning it must be looked upon as an erratic element (a "shock") which *itself* has taken care of the realistically existing bounds, or as an element given in some other way satisfying the bounds.

One very simple way in which we can explicitly take account of at least one important bound on the development of the economy, is by using the Harrod-Domar way of thinking where the need for *real capital* is introduced explicitly (and expressed through the capital-to-output ratio) and where this need for capital is connected with the *investment activity*. I do not think that the Harrod-Domar model gives an adequate analysis of the growth process in the economy.[6] In particular it is entirely inadequate for a discussion of the *decisions* connected with the growth problem in an underdeveloped country.[7] But the Harrod-Domar model takes at least account of one aspect of development which is completely neglected in the Hicksian model as described in (1)-(4). Another bound that may be considered is (148) below. The introduction of this latter bound will not restrict the number of degrees of freedom, so even with such a bound introduced, the model does not determine the actual solution of the system but will depend on what the autonomous investments H_t actually have been, subject to the constraint (148).

In the present paper I will, however, disregard this aspect of the Hicksian model and consider the nature of the solution as it follows from (1)-(5).

---

[6] See for instance my paper in *Econometrica*, July 1961.

[7] Some of my reasons for thinking so are given for instance in section III 1.a in my October 1960 paper in the journal *L'Egypt Contemporaine*, Cairo, and on pages 62–63 in my paper "Economic planning and the growth problem in developing countries", in the journal *Statsøkonomisk Tidsskrift*, Oslo, 1961.

THE SOLUTION BY RECURRENCE

It is essential to give the explicit solution of (5) *for any arbitrarily given function* H_t. We shall first approach this problem by the recurrence method.

Let the size of net national income in the points of time 0 and (-1), i.e. Y_0 Y_{-1}, be *given* as initial conditions. From (5) we then derive successively

$$\begin{aligned} Y_1 &= \mu Y_0 + \nu Y_{-1} + H_1, \\ Y_2 &= \mu^2(1+\omega) Y_0 + \mu\nu Y_{-1} + \mu H_1 + H_2, \\ Y_3 &= \mu^3(1+2\omega) Y_0 + \mu^2\nu(1+\omega) Y_{-1} + \mu^2(1+\omega) H_1 + \mu H_2 + H_3, \\ Y_4 &= \mu^4(1+3\omega+\omega^2) Y_0 + \mu^3\nu(1+2\omega) Y_{-1} + \mu^3(1+2\omega) H_1 \\ &\qquad + \mu^2(1+\omega) H_2 + \mu H_3 + H_4 \end{aligned} \qquad \dots \quad (8)$$

(Parikh has continued this recurrence up to $t = 15$. i.e. up to Y_{15}).

where
$$\omega = \frac{\nu}{\mu^2}. \qquad \dots \quad (9)$$

We can assume that
$$\mu \neq 0 \qquad \dots \quad (10)$$
so that the parameter ω defined by (9) exists.

It is readily seen that in general the coefficient of Y_0 in the explicit expression for Y_t will be μ^t times a polynomial in ω of the order Ent $t/2$ where

Ent q denotes the largest integer contained in q. ... (11)

Denoting the coefficients of this polynomial by a_{ti} we can write this polynomial

$$A_t = a_{t0} + a_{t1}\omega + a_{t2}\omega^2 + \dots \text{ up to the term with } \omega^{\text{Ent}\, t/2}. \qquad \dots \quad (12)$$

By carrying out the recurrence procedure (8) for the first terms it is easily seen that in general the first two coefficients are

$$a_{t0} = 1 \qquad (t = 1, 2, \dots, \infty) \qquad \dots \quad (13)$$

$$a_{t1} = t-1 \qquad (t = 1, 2, \dots, \infty). \qquad \dots \quad (14)$$

The general expression for the coefficients a_{ti} is then easily obtained by noticing that the general expression for Y_t must be linear in $Y_0, Y_{-1}, H_1, H_2, \dots, H_t$ and that the expression must hold good *identically* in these $t+2$ magnitudes.

Therefore if we put

$$Y_{-1} = H_1 = H_2 = \dots = H_t = 0, \qquad \dots \quad (15)$$

but
$$Y_0 \neq 0$$

and insert into (5) the expression for Y_t thus obtained, we get after having divided by Y_0 (assumed different from zero)

$$\mu^t \Sigma_i a_{ti} \omega^i = \mu . \mu^{t-1} \Sigma_i a_{t-1,\, i}\, \omega^i + \nu . \mu^{t-2} \Sigma_i a_{t-2,\, i}\, \omega^i. \qquad \dots \quad (16)$$

The last expression in the right member of (16) can be written $\mu^t \Sigma_i a_{t-2,\, i}\, \omega^{i+1} = \mu^t \Sigma_i a_{t-2,\, i-1} \omega^i$, so that (16) can also be written

$$\mu^t \Sigma_i (a_{ti} - a_{t-1,\, i} - a_{t-2,\, i-1}) \omega^i = 0 \qquad (t = 2, 3, \dots, \infty). \qquad \dots \quad (17)$$

By (10) this gives

$$\Sigma_i(a_{ti}-a_{t-1,\,i}-a_{t-2,\,i-1})\,\omega^i = 0 \qquad (t = 2, 3, \ldots, \infty). \qquad \ldots \quad (18)$$

Since the polynomial (18) is to be zero for any value of ω, its coefficients must be zero separately. Hence

$$a_{ti}-a_{t-1,i} = a_{t-2,\,i-1} \qquad (t = 2, 3, \ldots, \infty). \qquad \ldots \quad (19)$$

Adding the equations (19) for $t, t-1, \ldots 3$, we get

$$a_{ti} = a_{2i}+\sum_{\tau=3}^{t} a_{\tau-2,\,i-1} \qquad (t = 3, 4, \ldots, \infty). \qquad \ldots \quad (20)$$

In other words $$a_{ti} = a_{2i}+\sum_{\tau=1}^{t-2} a_{\tau,\,i-1} \qquad (t = 3, 4, \ldots, \infty). \qquad \ldots \quad (21)$$

We will only use this equation for $i = 2, 3, \ldots \infty$. Since by the remark before (11)

$$a_{2i} = 0 \qquad \text{for } i \geqslant 2 \qquad \ldots \quad (22)$$

we get from (21) $$a_{ti} = \sum_{\tau=1}^{t-2} a_{\tau,\,i-1} \qquad \begin{pmatrix} t = 3, 4, \ldots, \infty \\ i \geqslant 2 \end{pmatrix}. \qquad \ldots \quad (23)$$

This together with (13), (14) and the general formula for binomial coefficients

$$\sum_{\tau=0}^{p}\binom{\tau-i}{i}=\binom{p-i+1}{i+1} \qquad \begin{pmatrix} p \geqslant 0 \\ i \geqslant 0 \end{pmatrix} \qquad \ldots \quad (24)$$

we get $$a_{t2} = \sum_{\tau=1}^{t-2} a_{\tau1} = \sum_{\tau=1}^{t-2}\binom{\tau-1}{1}=\binom{t-2}{2} \qquad (t = 3, 4, \ldots, \infty). \qquad \ldots \quad (25)$$

From (8) we see that the last formula also holds good for $t = 1, 2$.

Hence we have quite generally

$$a_{t2} = \binom{t-2}{2} \qquad (t = 1, 2, \ldots \infty). \qquad \ldots \quad (26)$$

Further we get by (22), (23) and (25)

$$a_{t3} = \sum_{\tau=1}^{t-2} a_{\tau2} = \sum_{\tau=1}^{t-2}\binom{\tau-2}{2}=\binom{t-3}{3} \qquad (t = 3, 4, \ldots \infty). \qquad \ldots \quad (27)$$

Since by (8) this formula also holds good for $t = 1, 2$, we have quite generally

$$a_{t3} = \binom{t-3}{3} \qquad (t = 1, 2, \ldots, \infty). \qquad \ldots \quad (28)$$

Continuing in this way and taking account of (13) and (14) we get

$$a_{ti} = \binom{t-i}{i} \qquad \begin{pmatrix} t = 1, 2, \ldots, \infty \\ i = 0, 1, \ldots, \infty \end{pmatrix}. \qquad \ldots \quad (29)$$

We can check this formula by inserting into (5) the expression obtained for Y_t, Y_{t-1}, Y_{t-2} by using (29) and the assumptions (15). This gives for the left member, after dividing by Y_0

$$\mu^t\Sigma_i\binom{t-i}{i}\omega^i \qquad \ldots \quad (30)$$

and for the right member we get

$$\mu\cdot\mu^{t-1}\,\Sigma_i\binom{t-1-i}{i}\omega^i+\nu\cdot\mu^{t-2}\,\Sigma_i\binom{t-2-i}{i}\omega^i=\mu^t\Sigma_i\binom{t-1-i}{i}\omega^i+\mu^t\Sigma_i\binom{t-2-i}{i}\omega^{i+1}.$$

The last expression here is equal to $\mu^t\Sigma_i\binom{t-1-i}{i-1}\omega^i$. The total right member is consequently

$$\mu^t\Sigma_i\left[\binom{t-1-i}{i}+\binom{t-1-i}{i-1}\right]\omega^i. \qquad \ldots \quad (31)$$

By the classical formula for binomial coefficients

$$\binom{p}{i-1}+\binom{p}{i}=\binom{p+1}{i}\begin{pmatrix}p\geqslant 0\\ i\geqslant 0\end{pmatrix} \qquad \ldots \quad (32)$$

the bracket in (31) reduces to $\binom{t-i}{i}$ so that (31) is seen to be identical with (30).

From (8) we see heuristically that:

the coefficient of Y_{-1} *in* Y_t *is equal to* ν *times the coefficient of* Y_0 *in* Y_{t-1} ... (33)

and *the coefficient of* H_τ *in* Y_t *is equal to the coefficient of* Y_0 *in* $Y_{t-\tau}$. ... (34)

The two rules (33)-(34) can easily be verified by continuing the recurrence.

The general explicit expression for Y_t will consequently be

$$Y_t=\left[\mu^t\Sigma_i\binom{t-i}{i}\omega^i\right]Y_0+\left[\mu^{t-1}\nu\Sigma_i\binom{t-1-i}{i}\omega^i\right]Y_{-1}$$

$$+\sum_{\tau=1}^{t}\left[\mu^{t-\tau}\Sigma_i\binom{t-\tau-i}{i}\omega^i\right]H_\tau \quad (t=1,2,\ldots,\infty). \qquad \ldots \quad (35)$$

The brackets in (35) indicate the coefficients of Y_0, Y_{-1}, H_1, H_2, ..., H_T.

The first line in (35) can also be written

$$\mu^t\Sigma_i\frac{1}{i}\binom{t-i-1}{i-1}\left[\mu(t-i)Y_0+\nu(t-2i)Y_{-1}\right]\omega^i$$

but if it is desired to bring out explicitly the separate effect of each initial condition, (35) is the proper form to use.

By inserting the general expression (35) into (5) we can check that the equation holds good whatever values we attribute to μ, ω, Y_0, Y_{-1}, H_1, H_2, ..., H_t.

Indeed, inserting (35) into (5) and bringing all the coefficients of Y_0, Y_{-1}, $H_1, \ldots, H_t$ together on one side, we get

$$\begin{aligned}
&\left[\mu^t \left\{ \Sigma_i \binom{t-i}{i} \omega^i - \Sigma_i \binom{t-i-1}{i} \omega^i - \Sigma_i \binom{t-i-2}{i} \omega^{i+1} \right\}\right] Y_0 \\
&+ \left[\mu^{t-1}\nu \left\{ \Sigma_i \binom{t-i-1}{i} \omega^i - \Sigma_i \binom{t-i-2}{i} \omega^i - \Sigma_i \binom{t-i-3}{i} \omega^{i+1} \right\}\right] Y_{-1} \\
&+ \sum_{\tau=1}^{t-2} \left[\mu^{t-\tau} \left\{ \Sigma_i \binom{t-\tau-i}{i} \omega^i - \Sigma_i \binom{t-\tau-i-1}{i} \omega^i - \Sigma_i \binom{t-\tau-i-2}{i} \omega^{i+1} \right\}\right] H_\tau \\
&+ \left[\mu \left\{ \Sigma_i \binom{1-i}{i} \omega^i - \Sigma_i \binom{0-i}{i} \omega^i \right\}\right] H_{t-1} + [1-1] H_t.
\end{aligned}$$

In the last summation in first bracket above we replace i by $(i-1)$. By (32) the bracket is then seen to be zero. And the same applies to the following two brackets. The last two brackets are obviously zero.

THE SOLUTION THROUGH CHARACTERISTIC ROOTS

Another way to approach the solution of (5) is through characteristic roots. I repeat that it is essential to obtain the explicit solution of (5) for *any arbitrarily given function* H_t, not only for special analytical forms of H_t.

There are many text-books treating difference equations. But my experience is that in order to find in the text-books the results one needs, in a given problem, one has nearly always to wade through an enormous amount of symbolism. And the meaning of this symbolism can frequently be deciphered only by reading a great number of chapters different from the one where the results one is looking for is actually to be found. And even if one succeeds in deciphering the symbols the results are often not in *simple* form as one needs.

In view of the great importance for economic analysis of linear difference equations with constant coefficients and with an *arbitrary* additional term (such as H_t in (5)) I developed many years ago in my Oslo lecture (in Norwegian) an approach which is at the same time rigorous and, I believe, very easy to follow.[8] I shall reproduce here the "prescription" for the solution which this approach leads to. And I shall do it for the most general type of linear difference equation (with real constant coefficients), even though this most general case is more than needed in order to handle the simple equation (5).

The complete and the truncated difference equation. Let the difference equation in the unknown time function Y_t be

$$a_0 Y_t + a_1 Y_{t+1} + \ldots + a_n Y_{t+n} = H_{t+n} \qquad \ldots (36)$$

where $a_0, a_1, \ldots, a_n$ are given real coefficients and H_{t+n} an arbitrarily given real function of time. The integer n is called the *order* of the equation.[9]

---

[8] Appendix 7 in "Notater til ökonomisk teori" (mimeographed) 4. Edition, Oslo, 1947. Also, 'Ordinære lineære differentiallikninger", October, 1961.

[9] Several formulae pertaining to the solution of (36) will turn out in a much more handy form if we let a_ν denote the coefficient of $Y_{t+\nu}$ instead of the coefficient of

National Institute of Education

It does not restrict generality if we assume

$$a_n = 1. \qquad \ldots \quad (37)$$

Indeed if $a_n \neq 0$ we can simply divide the equation by a_n and if $a_n = 0$ *the equation* is simply an equation of the same kind but of lower order than n. In the sequel we will therefore assume (37).

The equation (36) with the additional term H_{t+n} will be called the *complete*, or non-homogeneous, equation.

To this complete, or non-homogeneous, equation we associate the corresponding *truncated*, or homogeneous, equation which is obtained by simply leaving out the additional term H_{t+n}, i.e. by putting $H_{t+n} = 0$ for all t. This truncated equation is consequently

$$a_0 Y_t + a_1 Y_{t+1} + \ldots + a_n Y_{t+n} = 0. \qquad \ldots \quad (38)$$

The solution of the truncated difference equation by means of characteristic roots. We will find first the general solution of the truncated equation (38). The basic fact in this connection (proved in an elementary way in my Oslo lecture) is the following: If we can only find in some way or other n *particular* solutions that are *linearly independent* over t, the problem is solved, because a linear combination of these n linearly independent particular solutions taken with arbitrary constant coefficients (constant in the sense of being independent of t) will give the general solution of (38). That is to say any function of t which satisfies (38) can be produced by attributing appropriate values to the n arbitrary constants.

Writing out in full: If $Y_{t1}, Y_{t2}, \ldots, Y_{tn}$ are n particular solutions which are *linearly independent* over t, then

$$Y_t = C_1 Y_{t1} + C_2 Y_{t2} + \ldots + C_n Y_{tn}, \qquad \ldots \quad (39)$$

where $C_1, C_2, \ldots, C_n$ are n arbitrary constants, will be the general solution of (38) in the sense that *any* function of t which is to satisfy (38) *must* be obtainable by attributing appropriate values to $C_1, C_2, \ldots, C_n$ in (39).

The fact that n function $Y_{t1}, Y_{t2}, \ldots, Y_{tn}$ are linearly independent over t means that it is *not* possible to find an *effective* set of constants $S_1, S_2, \ldots, S_n$ such that

$$S_1 Y_{t1} + S_2 Y_{t2} + \ldots + S_n Y_{tn} = 0, \quad \text{identically in } t. \qquad \ldots \quad (40)$$

The magnitudes $S_1, S_2, \ldots, S_n$ being *constant* means they are independent of t, and the set $S_1, S_2, \ldots, S_n$ being *effective* means at least one of these constants is different from zero.

The fact that the linear combination (39) with $C_1, C_2, \ldots, C_n$ as arbitrary coefficients will satisfy (38) if all the individual functions $Y_{t1}, Y_{t2}, \ldots, Y_{tn}$ satisfy (38), is obvious. This is immediately seen by inserting (39) into (38). And this conclusion obviously holds good whether the individual functions are linearly independent or not. The basic proposition in the theory of linear homogeneous difference equation is the *inverse* of this, namely, that *any* solution of (38) must be of the form (39) provided the particular solution $Y_{t1}, Y_{t2}, \ldots, Y_{tn}$ are linearly independent.

The problem of finding the general solution of (38) is therefore only to find *in some way or other* n linearly independent particular solutions of (38).

The general procedure for finding particular solutions of (38) is to try if an *exponential function*

$$Y_t^p = \lambda^t \ (p = \text{“particular”}) \qquad \ldots \ (41)$$

should be a particular solution, if the constant λ is chosen in an appropriate way.

Inserting (41) into (38) we see that if (41) is to be a solution of (38) we must have

$$a_0\lambda^t + a_1\lambda^{t+1} + \ldots + a_n\lambda^{t+n} = 0. \qquad \ldots \ (42)$$

This equation is to hold good for any value of t. Furthermore, since λ^t is different from zero if $\lambda \neq 0$ and t finite, we can divide the equation by λ^t. This shows that *if* (41) is to be a solution we *must* have

$$a_0 + a_1\lambda^1 + \ldots + a_n\lambda^n = 0. \qquad \ldots \ (43)$$

In other words *if* (41) is to be a solution, it is *necessary* that λ be a root of the n-th degree polynomial written in the middle member of (43).

The equation defined by (43) is called the *characteristic equation* for (38). Further, the polynomial $f(\lambda) = a_0 + a_1\lambda^1 + \ldots + a_n\lambda^n$ is called the *characteristic polynomial* for (38).

The characteristic equation (43) is simply an *algebraic equation* in λ with given coefficients. *This equation does not depend on t at all.* This is the essence of our tentative procedure for finding particular time functions that will satisfy (38).

The above reasoning shows that *if* (41) is to be a particular solution of (39) then λ *must* be a root of the characteristic equation (43). But the inverse also holds good. Obviously *any* root of the characteristic equation will yield a particular solution (41) which satisfies (38). Therefore if by any method whatsoever we have found a root of (43), we have by this fact also found a particular solution of (38).

It is a classical algebraic fact that a polynomial of degree n has exactly n zeros. This means that the characteristic equation (43) has exactly n roots (when each root is counted as many times as is indicated by its multiplicity). We may denote the roots $\lambda_1, \lambda_2, \ldots, \lambda_n$. Hence if the n roots of the characteristic equation (43) are all *different* (i.e. all of them of multiplicity 1), we have immediately solved the problem of finding the general solution of (38). Two exponential functions

$$Y_t^p = \lambda_p^t \quad \text{and} \quad Y_t^q = \lambda_q^t \qquad \ldots \ (44)$$

where $\lambda_p \neq \lambda_q$ are indeed always linearly independent over t.

The roots of (43) may be real or complex, but whatever they are, the above remark about linear independence always applies provided the roots are different.

If (43) has complex roots it is a classical fact that they must always occur in *conjugate pairs*, i.e. if

$$\lambda = \theta + i\delta \qquad i = \sqrt{-1} \qquad \dots \quad (45)$$

is a root with θ and δ real and $\delta \neq 0$ then also

$$\lambda_{\text{conj}} = \theta - i\delta \qquad \dots \quad (46)$$

must be a solution, and if $\delta \neq 0$ these two roots are different.

We are only interested in *real* solutions of (38). The only two problems that remain in order to find the general solution of (38) in the most general case where the coefficient $a_n, a_{n-1}, \dots, a_0$ of (38) are *any* given real numbers (with $a_n = 1$), are therefore, first to see if and in what way the two particular solutions that correspond to a complex root of (43) can be *combined* so as to give a *real* component of the general solution (a component that would absorb two different roots of (35)), and second to discuss the case of multiple roots of the equation.

The first of those two problems is a simple one, but the second demands a little closer analysis. It will turn out that in the case of multiple roots also we can find the necessary number n of *linearly independent* particular solutions. Therefore the complete discussion of all the n roots of the algebraic equation (43) (each root counted a number of times equal to its multiplicity) will yield the general solution of (38) in the general case where the coefficients $a_0, a_1, \dots, a_n$ are *any* given real numbers.

The problem of actually finding all the roots of (43) is a simple one for small values of n, say for $n = 2$. In this case it can even be done explicitly in terms of the coefficients. But for high values of n the problem cannot be handled except in the case where the coefficients are *numerically* given and effective numerical methods of computation are used.

If the roots are close together, the numerical task is always a difficult one. And if the roots are multiple a special numerical procedure must be applied. In the sequel I shall make a few remarks on the numerical aspect of finding the roots of the characteristic equation, but a full discussion of this numerical problem is beyond the scope of this paper.

For numerical reasons, both theoretical and practical, it is most convenient to concentrate first on the *real* roots. For each root λ_p which we succeed in determining, we can—by dividing the characteristic polynomial by $(\lambda - \lambda_p)$—*lower* the degree of the characteristic polynomial and thus reduce the difficulty of the search for further root.

If the problem is not so simple that the roots can be expressed explicitly in terms of the coefficients, then the problem must be approached numerically. The first thing is to draw a *graph* of $f(\lambda)$ as a function of λ. A study of this curve, including a study of the behaviour of the curve as λ tends towards $+\infty$ or $-\infty$, will indicate whether real roots are present, and if so, indicate approximately *where* they are. In other words we know already approximate values for the real roots. These values can subsequently be improved upon by known numerical methods. As a rule it is advantageous first to determine the root where the graph indicates that the curve cuts through the λ-axis in the *steepest* way. Here a numerical approximation method will work at its best. When this root, say λ_p, is determined with a high degree of accuracy, the polynomial $f(\lambda)$ is divided by $(\lambda-\lambda_p)$ and the work is continued on the resulting polynomial.

Already when considering the real roots we must be prepared for the possibility that any of the roots may be of higher multiplicity than one.

Since the polynomial $f(\lambda)$ is always continuous and with a continuous derivative, a graphical consideration is sufficient to indicate the existence of a root of multiplicity 2. The case of two roots coinciding means that the polynomial $f(\lambda)$ does not *cut through* the λ axis but only *touches* the λ axis in the point considered. This means that such a point λ_p must not only be a solution of $f(\lambda) = 0$, i.e. of (43) but must also be a solution of $f'(\lambda) = 0$, where $f'(\lambda) = df/d\lambda$ is the derivative of $f(\lambda)$ with respect to λ. This condition is not only necessary but also sufficient for the point λ_p to be of multiplicity 2, or higher. In general the necessary and sufficient condition for a point λ_p to be a root of multiplicity μ_p is that it satisfies at the same time all the μ_p equations

$$\begin{aligned} f(\lambda) &= 0 \\ f'(\lambda) &= 0 \\ f''(\lambda) &= 0 \qquad \ldots \quad (47) \\ &\cdots \\ f^{(\mu_p-1)}(\lambda) &= 0. \end{aligned}$$

In this case the polynomial $f(\lambda)$ is said to have contact of the order μ_p with the λ-axis on the point considered.

The curve that depicts the shape of $f(\lambda)$ will suggest if and where a multiple root is located. For instance, if it looks as if the curve has a minimum or a maximum on the λ-axis, this suggests the existence of a double root or more generally the existence of a root of even multiplicity i.e. $\mu_p = 2$ or 4 or 6. etc. If the curve has an inflexional point with the λ-axis itself as the inflexion tangent, this suggests the existence of a root of odd multiplicity $\mu_p = 3$ or 5 or 7 etc.

The equations (47) offer a grip on the problem of determining numerically and in an exact way the multiplicity of a root. We can analytically derive the successive polynomials $f'(\lambda), f''(\lambda) \ldots$ etc. and then for each of these polynomials draw a curve that depicts its course in the vicinity of the point considered. The first of them that has an ordinate different from zero in the point considered indicates the multiplicity of the root in question.

If the shapes of the curves are such that the situation in the vicinity of the point considered is not discernible in a sufficiently convincing way, one may have to draw successive sets of new curves on larger and larger scales but over a more and more restricted λ-range. This graphical procedure combined with one of the several available methods of numerical determination of the roots of non-linear equations will in practically all cases permit to determine all the real roots with any desired accuracy, as well as, the exact multiplicity of each root.

Any such real root, say λ_p, will yield a particular solution to (38) of the form (41). And if we have found such a particular solution, it is obvious that the function obtained by multiplying (41) by an arbitrary constant will also be a solution.

The case where simply a constant (independent of t) is a solution of (38) is only a special case of what has been said above. It means that the characteristic equation has the (simple or multiple) root $\lambda = 1$. Necessary and sufficient for this case is obviously that the sum of coefficients $a_0, a_1, \ldots, a_n$ is zero.

We know that if a certain number of roots $\lambda_p, \lambda_q, \ldots, \lambda_r$ are *different*, then the corresponding time-functions of the form (41) are not only solutions of (38) but they are also linearly independent. We have thus advanced at least part of the way towards a complete set of n linearly independent particular solutions. And if we actually have found n *different* roots of (38) the job is completed. It suffices then to form a linear aggregate with arbitrary constant coefficients of the n functions of the form (41). But if one or more of the roots are multiple we cannot complete the job merely by functions of the form (41). Indeed the number of roots of (42), when each root is counted a number of times according to its multiplicity, is exactly n. Hence we cannot get a sufficient number of linearly independent solutions merely by considering functions of the form (41). Therefore any multiple root raises a new problem of how to construct more linearly independent solutions.

To complete our discussion of real roots let λ_p be a real root of multiplicity μ_p where $\mu_p > 1$ (obviously we must have $\mu_p \leqslant n$). In this case any time-function of the form

$$Y_{t.p\mu} = t^{\mu}\lambda_p^t \qquad \ldots \quad (48)$$

(μ equal to any of the numbers $0, 1, \ldots, \mu_p-1$), must also satisfy the difference equation (38).

Indeed, by inserting (48) into (38) the left member becomes

$$\sum_{\nu=0}^{n} a_\nu(t+\nu)^{\mu}\lambda_p^{t+\nu} = \lambda_p^t\Big[\sum_{\nu=0}^{n} a_\nu(t+\nu)^{\mu}\lambda_p^{\nu}\Big]. \qquad \ldots \quad (49)$$

In the bracket to the right we develop $(t+\nu)^{\mu}$ by the binomial formula. This brings (49) into the form

$$\lambda_p^t \sum_{\nu=0}^{n} \sum_{i=0}^{\mu} a_\nu \binom{\mu}{i} \nu^i t^{\mu-i}\lambda_p^{\nu}. \qquad \ldots \quad (50)$$

The power z^i of any number z can be developed in terms of the factorials $z^{[j]}$ by the formula

$$z^i = \sum_{j=0}^{i} \binom{i}{j} B_{i-j}^{(-j)} z^{[j]} \qquad \dots \quad (51)$$

where the factorial $z^{[j]}$ is defined by

$$z^{[j]} = \begin{cases} 1 & \text{if } j = 0 \\ z(z-1)\dots(z-j+1) & \text{if } j \geqslant 1. \end{cases} \qquad \dots \quad (52)$$

and $B_{i-j}^{(-j)}$ are the Bernoullian numbers. They are pure numerical coefficients.[10]

Putting $z = \nu$ the expression (50) can therefore also be written

$$\lambda_p^{t-n} \sum_{\nu=0}^{n} \sum_{i=0}^{\mu} \sum_{j=0}^{i} a_\nu \binom{\mu}{i}\binom{i}{j} B_{i-j}^{(-j)} \nu^{[j]} t^{\mu-i} \lambda_p^{\nu}. \qquad \dots \quad (53)$$

In this expression the limits 0 and n of the first sum are *fixed* numbers. We can therefore move this summation sign to the right so that it becomes the last of the three summation operations to be performed. At the same time we replace λ_p^{ν} by $\lambda_p^{\nu-j}.\ \lambda^j$. This being done the expression (53) becomes

$$\lambda_p^{t-n} \sum_{i=0}^{\mu} \sum_{j=0}^{i} \binom{\mu}{i}\binom{i}{j} B_{i-j}^{(-j)}\, t^{\mu-i+j} \left[\sum_{\nu=0}^{n} a_\nu\, \nu^{[j]} \lambda_p^{\nu-j} \right]. \qquad \dots \quad (54)$$

The bracket to the right in this expression is the value in the point $\lambda = \lambda_p$ of the j-th derivative $f^{(j)}(\lambda)$ of the characteristic polynomial. Since we have assumed that μ is *less* than the multiplicity μ_p, and by the first two summation signs in (54) $i \leqslant \mu$ and $j \leqslant i$, and consequently $j \leqslant \mu$ we must have j less than the multiplicity μ_p. Consequently by (47) all the brackets in (54) must vanish. Hence the expression (54) must vanish for any value of t. Consequently, the left member of (49)—which is the same as (54)—must vanish identically in t. That is to say the function (48) must satisfy the difference equation (38).

This shows that if λ_p is a root of multiplicity μ_p of the characteristic equation (43), all the μ_p time-functions

$$\lambda_p^t,\ t\lambda_p^t,\ t^2\lambda_p^t,\ \dots\ t^{\mu_p-1}\lambda_p^t \qquad \dots \quad (55)$$

must satisfy (38). *All the μ_p time-functions are linearly independent* (and, of course, also linearly independent of corresponding time-functions corresponding to any other root λ_q).

Therefore to each root λ_p in (38), whether simple or multiple, we have determined as many linearly independent particular solutions of (38) as is indicated by the multiplicity of the root λ_p.

---

[10] See for instance formula (9b) and table 1, p. 12, in my memoire "Sur les semi-invariants et moments employes dans l'etude des distributions statistique." Det Norske Videnskap-Akademi i Oslo, 1926.

In the above argument we reasoned as if λ_p should be a *real* root. But the whole argument *is equally valid in the case of a complex root.* Therefore whatever the nature of the characteristic polynomial we have actually determined exactly n linearly independent solutions of the difference equation (38). Hence we have obtained the *general* solution of this equation.

This general solution will be of the form (39), i.e. it will contain n arbitrary constants.

Furthermore, these constants appear in the formula in such a way that if we consider the equation (39) for n arbitrarily given points of time where the values of Y_t are *prescribed*, and if these points of time are such that the $n \times n$ matrix of values Y_{ti} in the right member of (39) is non-singular, we can determine — and in a unique way — the values which it is necessary and sufficient to attribute to the constants in order that the solution shall assume the prescribed n values of Y_t.

Therefore the only thing which now remains is to show how the complex solutions can be combined into pairs in such a way that we obtain *real* time-functions and we retain *the same number n* of arbitrary constants and these occur in such a manner that by a suitable choice of them we can make the solution to assume n prescribed values of Y_t.

To show how this combination of complex roots can be achieved let us consider a complex root (45). The fact that (45) — and hence (46) — satisfies the characteristic equation means that

$$\sum_{\nu=0}^{n} a_\nu \ (\theta + i\delta)^\nu = 0. \qquad \ldots \quad (56)$$

Developing the parenthesis in (56) by the binomial formula we get

$$\sum_{\nu=0}^{n} a_\nu \sum_{\mu=0}^{\nu} \binom{\nu}{\mu} \theta^{\nu-\mu} \, (i\delta)^\mu. \qquad \ldots \quad (57)$$

The summation over μ can be split up in the following four sums

$$\left[\sum_{\mu=0,4,8..} \binom{\nu}{\mu} \theta^{\nu-\mu} \, \delta^\mu - \sum_{\mu=2,6,10\ldots} \binom{\nu}{\mu} \theta^{\nu-\mu} \, \delta^\mu \right]$$

$$+ i \left[\sum_{\mu=1,5,9\ldots} \binom{\nu}{\mu} \theta^{\nu-\mu} \, \delta^\mu - \sum_{\mu=3,7,11\ldots} \binom{\nu}{\mu} \theta^{\nu-\mu} \, \delta^\mu \right] \qquad \ldots \quad (58)$$

where all the summations over μ are to be continued to the last not vanishing term.

The first bracket in (58) is the real part and the second bracket the imaginary part. Since these parts must be zero separately in order that the whole expression be zero, we get

$$\sum_{\nu=0}^{n} a_\nu \left[\sum_{\mu=0,4,8\ldots} \binom{\nu}{\mu} \theta^{\nu-\mu} \, \delta^\mu - \sum_{\mu=2,6,10\ldots} \binom{\nu}{\mu} \, \theta^{\nu-\mu} \delta^\mu \right] = 0 \qquad \ldots \quad (59)$$

$$\sum_{\nu=0}^{n} a_\nu \left[\sum_{\mu=1,5,9\ldots} \binom{\nu}{\mu} \theta^{\nu-\mu} \, \delta^\mu - \sum_{\mu=3,7,11\ldots} \binom{\nu}{\mu} \theta^{\nu-\mu} \, \delta^\mu \right] = 0. \qquad \ldots \quad (60)$$

These two equations contain only real numbers and are therefore amendable to numerical computation. The values of δ and θ that satisfy these two simultaneous equations define a complex root of the form (45).

It is easily seen that if the set of two numbers (δ, θ) satisfy (59)-(60), the conjugate set $(-\delta, \theta)$ must also satisfy. These two roots $(\theta+i\delta)$ and its conjugate $(\theta-i\delta)$ are precisely the roots we want to combine.

The numerical solution of the system consisting of the two simultaneous equations (59)-(60) in the two real variables δ and θ will, of course, as a rule be more laborious than the solution of the single equation (43) in one real variable λ. But I shall not go into details about this. I shall only show in a subsequent note that the numerical work can also, if we like, be performed in terms of real trigonometric functions.

In order to show that a complex root of the characteristic equation (43) and its conjugate root will together yield two linearly independent *real* time-functions satisfying (38), we make the transformation[11]

$$\lambda = e^{\gamma} \qquad \dots \quad (61)$$

$e = 2.71828\dots$ (basis of natural logarithms)

where[12] $$\gamma = \beta+i\alpha \quad i = \sqrt{-1},\ \beta \text{ and } \alpha \text{ real numbers} \qquad \dots \quad (62)$$

and the complex form of λ is given by (45). Then by classical formulae

$$\lambda = \theta+i\delta = r(\cos\alpha+i\sin\alpha) \qquad \dots \quad (63)$$

where $$r = \sqrt{\theta^2+\delta^2} \quad \theta = r\cos\alpha; \quad \delta = r\sin\alpha. \qquad \dots \quad (64)$$

All the magnitudes in (64) are real, and $r = 0$. Obviously it does not restrict generality if we assume

$$0 \leqslant \alpha < 2\pi. \qquad \dots \quad (65)$$

Raising the *left* member of (61) to the power t is the same as to raise the last member of (63) to the power t. By classical formulae this gives

$$\lambda^t = r^t\ (\cos\alpha t+i\sin\alpha t). \qquad \dots \quad (66)$$

On the other hand by raising directly the *right* member of (61) to the power t we get

$$e^{\gamma t} = e^{(\beta+i\alpha)t} \qquad \dots \quad (67)$$

which by classical formulae reduces to

$$e^{\gamma t} = e^{\beta t(\cos\alpha t+i\sin\alpha t)}. \qquad \dots \quad (68)$$

Comparing (68) with (66) we see that

$$r = e^{\beta}. \qquad \dots \quad (69)$$

Therefore (67) can also be written

$$e^{\gamma t} = r^t e^{i\alpha t}. \qquad \dots \quad (70)$$

---

[11] When handling differential equations as distinct from difference equations we are led immediately to a characteristic equation in λ not to one in γ and hence need not consider a subsequent transformation of the kind (61).

[12] The letters α and β in (62) should not be confused with the structural constants α and β in (1)-(7).

The conjugate root is obtained by replacing δ by $(-\delta)$, which by the first and last equation in (64) is the same as to replace α by $(-\alpha)$ but retaining r. In other words the two time-functions

$$r^t e^{i\alpha t} \quad \text{and} \quad r^t e^{-i\alpha t} \qquad \dots \quad (71)$$

are both solutions of (38). They are linearly independent when $\alpha \neq 0$, which simply means that the root in question is actually complex.

To combine the two time-functions in (71) in such a way as to obtain two *real* and linearly independent time-functions we multiply the first of them by an arbitrary complex constant and the second of them by a complex constant that is the conjugate of the first.

Let the two conjugate constants be

$$(A-iB) \quad \text{and} \quad (A+iB) \quad i = \sqrt{-1} \qquad \dots \quad (72)$$

where A and B are arbitrary *real* constants. Using the classical formula

$$e^{iz} = \cos z + i \sin z \quad i = \sqrt{-1} \qquad \dots \quad (73)$$

we get

$$(A-iB)r^t e^{i\alpha t} + (A+iB)r^t e^{-i\alpha t} = 2Ar^t \cos \alpha t + 2Br^t \sin \alpha t. \qquad \dots \quad (74)$$

The two time-functions written to the right in (74) are real and linearly independent. The aggregate satisfies the difference equation for any values we may choose for A and B. Choosing first $A = \frac{1}{2}$, $B = 0$ and next $A = 0$, $B = \frac{1}{2}$ we see that each of the two functions

$$r^t \cos \alpha t \quad \text{and} \quad r^t \sin \alpha t \qquad \dots \quad (75)$$

is a solution of (38). They are linearly independent if $r \neq 0$ and $\alpha \neq 0$. These two conditions are obviously fulfilled if we actually have a complex root.

If we like we can combine the two time-functions (75)—each taken with an arbitrary real coefficient—in such a way that the aggregate appears in the form of *a wave component* satisfying (38). The wave will be damped, explosive or non-changing according to the value of r in the following way

$$r < 1 \text{ damped} \quad r > 1 \text{ explosive,} \quad r = 1 \text{ non-changing.} \qquad \dots \quad (76)$$

We have indeed[13]

$$Ar^t \cos \alpha t + Br^t \sin \alpha t = Kr^t \sin (v + \alpha t) \qquad \dots \quad (77)$$

where

$$K = |\sqrt{A^2+B^2}| \; ; \; A = K \sin v \; ; \; B = K \cos v. \qquad \dots \quad (78)$$

Obviously it does not restrict generality if we assume

$$0 \leqslant v \leqslant 2\pi \qquad \dots \quad (79)$$

K is the *scale factor* and v *the phase* of the wave component. Between the two sets of constants (A, B) in the left member of (77) and (K, v) in the right member there is a one-to-one correspondence. Any given set (A, B) will by (78)-(79) lead to a uniquely determined set (K, v). And inversely any given set (K, v) will by the last two equations in (78) lead to a uniquely determined set (A, B). Thus, whether we look at the left or the right expression in (77) we have two arbitrary constants at disposal.

---

[13] In (77) A and B are used in the sense of one half the values A and B used in (73). This, of course is only a typographical convenience.

For computational reasons it is as a rule simplest to use the left hand expression in (77). Here the nature of the aggregate as the sum of two linearly independent time-functions is also brought clear to light. On the other hand if we want to plot a graph of the aggregate as a time-function and study its nature as a *wave component* satisfying (38), the right hand expression in (77) is useful.

The treatment of *multiple* complex roots of (43) can be handled in a way which is quite analogous to the way in which we handled multiple *real* roots in (47)-(55).

Note on the solution of the characteristic equation by means of trigonometric functions. Inserting (61)-(63) into the left member of the characteristic equation (34) we get by (73)

$$\sum_{\nu=0}^{n} a_\nu \lambda^\nu = \sum_{\nu=0}^{n} a_\nu r^\nu (\cos \alpha\nu + i \sin \alpha\nu) \tag{80}$$

separating here the real and imaginary parts we see that the single equation (43) in the complex variable λ is equivalent with the following two equations in the real magnitudes r and α

$$\sum_{\nu=0}^{n} a_\nu r^\nu \sin \alpha\nu = 0 \quad \text{and} \quad \sum_{\nu=0}^{n} a_\nu r^\nu \cos \alpha\nu = 0 \qquad \ldots \quad (81)$$

where it does not restrict generality if we impose (65).

If we want to handle the two simultaneous characteristic equations numerically in the trigonometric form (81) the following formula may be found useful. Putting for brevity

$$s = \sin \alpha \qquad c = \cos \alpha \qquad \ldots \quad (82)$$

we have

$$\begin{aligned} \sin 2\alpha &= 2cs & \cos 2\alpha &= 1-2s^2 \\ \sin 3\alpha &= s(3-4s^2) & \cos 3\alpha &= c(1-4s^2) \\ \sin 4\alpha &= cs(4-8s^2) & \cos 4\alpha &= 1-8s^2+8s^4. \end{aligned} \qquad \ldots \quad (83)$$

These formulae can easily be extended to subsequent values of ν by the recurrence formula

$$\sin(\nu+1)\alpha = c\,.\sin \nu\alpha + s.\cos \nu\alpha \quad \text{and} \quad \cos(\nu+1)\alpha = c.\cos \nu\alpha - s.\sin \nu\alpha \quad \ldots \quad (84)$$

The general solution of the complete (non-truncated equation). Let Y_t^P and Y_t^Q be any two particular solutions of the complete (non-homogeneous) difference equation (36). I have used capital letters for superscripts to indicate that the two time-functions are solutions of the complete equation.

If Y_t^P and Y_t^Q are any solution of the complete equation, then their difference $(Y_t^P - Y_t^Q)$ must be a solution of the corresponding *truncated* (homogeneous) equation. This is immediately seen by inserting first Y_t^P and next Y_t^Q in the complete equation and taking the difference between the two equations thus obtained.

This being so, let Y_t^Q stand for any particular solution (of the complete equation) which we *know* in its whole course over t, and let Y_t^P stand for a particular solution (of the complete equation) which is so far not known in its whole course over t but is only fixed through its specialized initial conditions. Then there must, by the argument in the beginning of this section, exist *some specific* solution y_t^P of the *truncated* equation which represents the difference $(Y_t^P - Y_t^Q)$. In other words, for any t we must have

$$Y_t^P = y_t^P + Y_t^Q \qquad \dots \quad (85)$$

where y_t^P is a particular solution of the truncated equation, its particularization depending on what particular solutions Y_t^P and Y_t^Q we are considering.

Inversely if in the right member of (85) we insert any particular solution y_t^P (of the truncated equation) which we may happen to think of, we will through (85) have generated some particular solution Y_t^P of the complete equation. Furthermore, if we let y_t^P vary in all possible ways, we will by (85) generate a whole lot of different time-functions y_t^P, all of which are solutions of the complete equation. In fact we will generate *all possible* solutions of the complete equation. This follows from the fact that "all possible" solutions y_t^P in the right member of (85) means a function with n arbitrary constants (they occur as we know linearly). And n arbitrary constants in the right member of (85) means that the function Y_t^P defined by (85) will contain n arbitrary constants (appearing linearly).

This shows that if we insert in (85) for y_t^P the *general* solution $Y_t^{\text{gen.trun}}$ of the truncated equation (this general solution will contain n arbitrary constants) we will have generated the *general* solution of the complete (non-homogeneous) equation. In other words

$$Y_t^{\text{gen.com}} = Y_t^{\text{gen.trun}} + Y_t^Q \qquad \dots \quad (86)$$

where $Y_t^{\text{gen.comp}}$ denotes the general solution of the complete (non-homogeneous) difference equation, and Y_t^Q denotes *any* particular solution of the complete equation which we have been able to obtain in some way or the other.

The formula (86) may give rise to many *different forms* for presenting the general solution of the complete equation. If we have found one particular solution Y_t^Q of the complete equation we can add to it any particular solution Y_t^q of the truncated equation. The result $Y_t^R = Y_t^Q + Y_t^q$ will also be a particular solution of the complete equation and can be used as the second term in the right member of (86).

The problem of finding the general solution of the complete equation is by (86) reduced to that of first finding the *general* solution of the *truncated* equation — a problem which was entirely solved in previous sections — and second finding *any* particular solution of the complete equation.

Sometimes we may be able to find a particular solution of the complete equation by a happy guess. But in general we cannot, of course, rely on this. We must therefore have a systematic way of finding one such particular solution.

The determination of a particular solution of the complete equation. The following is an elementary but rigorous procedure for obtaining a particular solution of the complete equation. The procedure is built on the knowledge of the *general* solution of the truncated equation.

Let the general solution of the truncated equation be known and let η_t be the particular solution of the truncated equation which is obtained by specifying the following initial conditions

$$\begin{aligned} a_n\eta_0 &= 1 \\ a_n\eta_1+a_{n-1}\eta_0 &= 0 \\ a_n\eta_2+a_{n-1}\eta_1+a_{n-2}\eta_0 &= 0 \\ \ldots \quad \ldots \quad \ldots & \\ a_n\eta_{n-1}+a_{n-1}\eta_{n-2}\cdots+a_1\eta_0 &= 0\,. \end{aligned} \qquad \ldots \quad (87)$$

The arguments 0, 1, 2, ... etc. on η in (87) are not obtained by a conventional choice of origin for t in η_t, but have the *absolute* meaning of 0, 1, 2, ... etc. They represent number of intervals of the same sort as occur in the difference equation itself.

More briefly (87) can be written

$$\begin{aligned} a_n\eta_0 &= 1 \\ \sum_{\nu=\tau}^{n} a_\nu\eta_{\nu-\tau} &= 0, \text{ for } \quad \tau = 1, 2, \ldots (n-1). \end{aligned} \qquad \ldots \quad (88)$$

This particular solution of the truncated equation will be called the cumulator.

Since by (37) $a_n \neq 0$, the first equation above determines η_0, similarly η_1 is determined by the second equation (if a_{n-1} should happen to be zero we get $\eta_1 = 0$). Similarly η_2 is determined by the third equation etc., and finally η_{n-1} by the n-th equation. These initial conditions $\eta_0, \eta_1, \ldots, \eta_{n-1}$ determine the particular solution in a unique way. The particular solution η_t obtained in this way will by definition satisfy the difference equation

$$\sum_{\nu=0}^{n} a_\nu\eta_{t+\nu} = 0 \qquad \text{for any value of } t. \qquad \ldots \quad (89)$$

Consider the time-function

$$Y_t^Q = \sum_{\tau=0}^{t-s} \eta_t H_{t-\tau} \;(t = s, s+1, \ldots, \infty) \qquad \ldots \quad (90)$$

where s is the earliest point of time such that H_T is known for $T = s, s+1, \ldots, \infty$. This means that $t = s$ is the earliest point of time for which the expression (90) is applicable. For precision of thinking it is essential to specify the point of time s. Since we may, of course, if we like *disregard* our knowledge of H_t over certain time ranges, the parameter s in (90) may be interpreted as *any* point of time which is such that in this point of time and in all *subsequent* points of time H_t is known. In this sense the upper limit of summation in (90) contains an *arbitrary* element.

The time-function (90) will satisfy the difference equation (36) for any point of time $t = s, s+1, s+2$ ... ad. infin.

Indeed, if we insert (90) into the left member of (36) we get

$$\sum_{\nu=0}^{n} a_\nu \, Y_{t+\nu}^Q = \sum_{\nu=0}^{n} a_\nu \sum_{\tau=0}^{t+\nu-s} \eta_t H_{t+\nu-\tau} = \sum_{\nu=0}^{n} \sum_{\tau=-\nu}^{t-s} a_\nu\eta_{\tau+\nu}H_{t-\tau} \quad (t = s, s+1, \ldots, \infty) \ \ldots \quad (91)$$

In the expression to the right in (91) we interchange the order of the two summation operations. Since

$$\sum_{\nu=0}^{n} \sum_{\tau=-\nu}^{t-s} = \sum_{\tau=0}^{t-s} \sum_{\nu=0}^{n} + \sum_{\tau=-1}^{-n} \sum_{\nu=-\tau}^{n} \qquad \ldots \quad (92)$$

the expression to the right in (91) becomes

$$\sum_{\tau=0}^{t-s} H_{t-\tau}\left[\sum_{\nu=0}^{n} a_\nu \eta_{\tau+\nu} \right] + \sum_{\tau=-1}^{-n} H_{t-\tau}\left[\sum_{\nu=-\tau}^{n} a_\nu \eta_{\tau+\nu} \right]. \qquad \ldots \quad (93)$$

The bracket in the first term of (93) is by (89) zero for any value of τ. Therefore the first term of (93) vanishes.

If we introduce $(-\tau)$ instead of τ as a summation affix in the last term of (93), this last term becomes

$$\sum_{\tau=1}^{n} H_{t+\tau}\left[\sum_{\nu=\tau}^{n} a_\nu \eta_{\nu-\tau} \right]. \qquad \ldots \quad (94)$$

For $\tau = n$ the bracket in (94) becomes $\sum\limits_{\nu=n}^{n} a_\nu \eta_{\nu-n} = a_n \eta_0$ which by the first equation in (88) is equal to 1. For $\tau = n$ (94) will therefore give H_{t+n}. For $\tau = 1, 2, \ldots (n-1)$ the bracket in (94) will vanish by the last equation in (88).

All that is left of (93)—which is the same as the expression in the left member of (91)—is consequently H_{t+n}. In other words the time-function $Y^{Q}_{t+\nu}$ satisfies the complete difference equation (36).

The *general* solution $Y_t^{\text{gen. comp}}$ of the *complete* difference equation is therefore by (68) and (90)

$$Y_t^{\text{gen. comp}} = Y_t^{\text{gen. trun}} + \sum_{\tau=0}^{t-s} \eta_\tau H_{t-\tau} \quad (t = s, s+1, \ldots, \infty) \qquad \ldots \quad (95)$$

$Y_t^{\text{gen. trun}}$ in (95) denoting the *general* solution of the truncated (homogenous) difference equation (38),—a solution that contains n arbitrary constants, which may, if we like, be chosen as real constants—, and η_τ in (95) denoting the *particular* solution of the truncated (homogeneous) difference equation (38) which satisfies the initial conditions (87).

Since the general solution of the truncated (homogeneous) difference equation (38) is given by the reasoning in (41)-(47), the right hand member of (95) gives the explicit expression for the general solution of the complete (non-homogeneous) difference equation (36). This general solution is applicable for $t = s, s+1 \ldots, \infty$ in the sense that (36) for $t = s, s+1 \ldots, \infty$ will be satisfied if we insert (95). This limitation on the applicability of (95) will, of course, not prevent us from assuming that the difference equation (36) *itself* is valid even for earlier points of time, but then the difference equation can only be used *by recurrence* for these earlier points.

Application to the Hicksian difference equation. In order to apply (95) to the equation (5) we replace in (5) t by $(t+2)$ and write the equation in the following form which is immediately comparable with (5)

$$-\nu Y_t - \mu Y_{t+1} + Y_{t+2} = H_{t+2}. \quad \dots \quad (96)$$

That is we have $$a_0 = -\nu, \quad a_1 = -\mu, \quad a_2 = 1. \quad \dots \quad (97)$$

The special case $\nu = 0$ means a first order equation. ... (98)

The special case $\mu = 0$ ... (99)

can be reduced to a first order equation by doubling of the interval considered. These first order cases will be discussed later. We begin by assuming

Both conditions $\nu \neq 0$ *and* $\mu \neq 0$ *fulfilled so that*
(96) *is a true second order difference equation.* ... (100)

In the case (100) we must distinguish between the three subcases,

$$\left.\begin{array}{ll} \text{I.} & \textit{two real and different roots of} \\ & \textit{the characteristic equation;} \\ \text{II.} & \textit{two real and equal roots;} \\ \text{III.} & \textit{two complex (and then necessarily} \\ & \textit{conjugate) roots: the wave-case;} \end{array}\right\} \quad \dots \quad (101)$$

We will discuss these cases one by one.

The roots of the characteristic equation corresponding to the truncated Hicksian difference equation. The characteristic equation is

$$\lambda^2 - \mu\lambda = \nu \quad \dots \quad (102)$$

the roots of which are $$\lambda_1 = \frac{\mu}{2}(1+\sqrt{1+4\omega}) \quad \dots \quad (103)$$

$$\lambda_2 = \frac{\mu}{2}(1-\sqrt{1+4\omega}). \quad \dots \quad (104)$$

By convention we interpret the square root sign in (103)-(104) as the positive branch, i.e.

$$\sqrt{1+4\omega} = \begin{cases} \text{absolute value of } \sqrt{1+4\omega} & \text{if } 1+4\omega \geqslant 0 \\ \sqrt{-1} \text{ times absolute value of } \sqrt{(1+4\omega)} & \text{if } 1+4\omega < 0. \end{cases} \quad \dots \quad (105)$$

The three cases (101) will be characterized by

I. if $1+4\omega > 0$ (Two real and different roots)
II. if $1+4\omega = 0$ (Two real and equal roots) ... (106)
III. if $1+4\omega < 0$ (Two complex roots).

In the case I: Two real and different roots we have by the convention (105)

absolute value of $\lambda_1 >$ absolute value of λ_2. ... (107)

In terms of the original parameters α, β, κ, λ of the model (1)-(4)[14] the nature of the characteristic equation may be characterised by the set of two parameters (μ, ω) where μ is defined by (6) and ω by

$$\omega = \frac{\lambda-\beta}{(\alpha+\beta+\kappa-\lambda)^2}. \qquad \dots \quad (108)$$

Or it may be characterised by the set of two coefficients (ν, π) where ν is defined by (7) and π by

$$\pi = \kappa+\alpha. \qquad \dots \quad (109)$$

Between the two sets (μ, ω) and (ν, π) there is a one-to-one correspondence, because we have

$$\mu = \pi-\nu. \qquad \dots \quad (110)$$

$$\omega = \frac{\nu}{(\pi-\nu)^2}. \qquad \dots \quad (111)$$

and

$$\nu = \omega\mu^2 \qquad \dots \quad (112)$$

$$\pi = \mu(1+\omega\mu). \qquad \dots \quad (113)$$

From the viewpoint of the *government strategy* in matters of current account expenditure it might be most convenient to consider the set (ν, π) since by (7) ν differs from the strategy parameter λ only by the addition of a *structurally* defined constant, namely $(-\beta)$, and by (109) π differs from the strategy parameter κ only by the addition of a *structurally* defined constant, namely α : But from the viewpoint of the explicit expression (35), obtained by recurrence, and from the viewpoint of the expression (103)-(104) for the characteristic roots the set (μ, ω) is unquestionably the most convenient.

The condition for the non-wave cases I and II which should be aimed at when determining government strategy—is simply

$$1+4\omega \geqslant 0. \qquad \dots \quad (114)$$

This can be looked upon as a (μ, ω) formulation of the non-wave condition. The corresponding (ν, π) formulation (in the cases of a true second order difference equation) is

$$(\pi-\nu)^2+4\nu \geqslant 0. \qquad \dots \quad (115)$$

In the special case where the difference equation degenerates to a first order difference equation—for instance because government strategy has the features $\lambda = \beta$ —the question of a wave formed evolution over time will, of course, not arise.

---

[14] The parameter λ occurring in the model must, of course, not be confounded with the unknown λ occurring in the characteristic equation (102). In the sequel no confusion is possible since we use the notation λ_1 and λ_2 for the two roots of the characteristic equation. We might perhaps have used a different letter for the parameter λ occurring in (5). The structural constants α and β in (1)-(7) should not be confused with the letters α and β in (62).

In the study of the true second order case of (96) the following formulae, which easily follow from (103)-(104) may be useful

$$2r^2 = \mu^2(1+2\omega) \qquad \ldots \quad (116)$$

$$\lambda_1+\lambda_2 = \mu \qquad \ldots \quad (117)$$

$$\lambda_1\lambda_2 = -\nu \qquad \ldots \quad (118)$$

$$\lambda_1-\lambda_2 = \mu\sqrt{1+4\omega} \qquad \ldots \quad (119)$$

$$\frac{\lambda_1^2-\lambda_2^2}{\lambda_1-\lambda_2} = \mu \qquad \ldots \quad (120)$$

$$\frac{\lambda_1^3-\lambda_2^3}{\lambda_1-\lambda_2} = \lambda_1^2+\lambda_1\lambda_2+\lambda_2^2 = \mu^2+\nu = \mu^2(1+\omega). \qquad \ldots \quad (121)$$

In the sequel we will simply disregard the wave-case, i.e., case III in (101) and (106) because this case *can be prevented so easily*, simply by imposing on the government strategy parameters κ, λ the condition (114), which is equivalent to (115). But then, of course, this condition must not be forgotten in any possible programming formulation.

Case I: *The explicit expression for the general solution of the complete Hicksian difference equation in terms of the characteristic roots when the roots are real and different.*

The cumulator η_t—cfr. (87), (88) and (89)—is in the present case

$$\eta_t = c_1\lambda_1^t+c_2\lambda_2^t \qquad \ldots \quad (122)$$

where the constant c_1 and c_2 ought to be determined so that the cumulator satisfies the initial conditions

$$\begin{aligned} a_2\eta_0 &= 1 \\ a_2\eta_1+a_1\eta_0 &= 0 \end{aligned} \qquad \ldots \quad (123)$$

where now $a_2 = 1 \quad a_1 = -\mu$ ... (124)

which gives $\eta_0 = 1 \quad \eta_1 = \mu.$ ... (125)

This gives, in the case $\lambda_1 \neq \lambda_2$

$$c_1 = \frac{\mu-\lambda_2}{\lambda_1-\lambda_2} \quad c_2 = \frac{\lambda_1-\mu}{\lambda_1-\lambda_2}. \qquad \ldots \quad (126)$$

Inserting this into (122) we get after some reductions

$$\eta_t = \frac{\lambda_1^{t+1}-\lambda_2^{t+1}}{\lambda_1-\lambda_2} \quad (t = \text{any point of time}). \qquad \ldots \quad (127)$$

The general solution of the complete difference equation (96) will therefore, (cf. (95)) be of the form

$$Y_t = C_1\lambda_1^t+C_2\lambda_2^t+\sum_{\tau=0}^{t-s}\eta_\tau H_{t-\tau} \quad (t = s, s+1 \ldots \infty) \qquad \ldots \quad (128)$$

s being a point of time such that H_t is known for $t = s$ and for all *subsequent* points of the time, and C_1 and C_2 being two arbitrary constants.

If we take—as we did in (35)—Y_0 and Y_{-1} as expressions for the initial conditions, and assume that H_t is known for $t = s, s+1 \ldots, \infty$ the procedure for expressing the coefficients C_1 and C_2 in terms of Y_0 and Y_{-1} will be, first to compute Y_1 and Y_2 by recurrence in terms of Y_0 and Y_{-1}, H_1 and H_2 (the result is given by the first two expressions in (8)) and then to put down the two equations which state that these expressions for Y_1 and Y_2 ought to be the same as the right member of (128).

Doing this, we get after some reduction, using (117)-(119) the following expression for the general solution of the complete equation (96)

$$Y_t = \eta_t Y_0 + \nu \eta_{t-1} Y_{-1} + \sum_{\tau=0}^{t-s} \eta_\tau H_{t-\tau} \qquad \ldots \quad (129)$$

$$(t = s, s+1 \ldots, \infty); \quad (s = 1, 2, \ldots, \infty)$$

By following the above argument literally we would have to put $s = 1$ in (129), but it is easily seen that (129) applies[15] for any $1 \leqslant s \leqslant t$. The parameter s is precisely the *arbitrary* element in the summation limit which we considered in connection with (90). By inserting (129) into the difference equation we see in fact that all values of H_T will disappear for $T > t$ where t is the earliest point of time that occurs for H_T when (129) is inserted in the difference equation. In other words if we only think of the use of (129) *in the difference equation*, we may attribute to H_T for $T < t$ any values we like (but it must be the *same* values of H_T for *all* the t in (129) that are used in the difference equation). On the other hand if we want to take (129) as an *independent definition* of a function Y_t regardless of the use we may want to make of this function, we must assume that H_T is known for $T \geqslant s$.

The formula (129) shows the basic role played by the cumulator η_t, defined as the particular solution of the truncated equation which satisfies the initial condition (123), or in general (87). *The whole problem is really solved as soon as the cumulator is obtained and the H_T are known for all the T we need.*

In (35) we assume that H_1 was the earliest known value of H_T. The same assumption may also be made in (129), i.e. we may put

$$s = 1. \qquad \ldots \quad (130)$$

This will reduce (129) to (35) identically in t and identically in the parameters μ and ω.

To permit an easier comparison with (35) we can transform the expression for the cumulator as follows, (cf. (119))

$$\eta_t = \frac{\lambda_1^{t+1} - \lambda_2^{t+1}}{\lambda_1 - \lambda_2} = \frac{1}{\varepsilon\mu}\left(\frac{\mu}{2}\right)^{t+1}[(1+\varepsilon)^{t+1} - (1-\varepsilon)^{t+1}] \qquad \ldots \quad (131)$$

where we have put for brevity

$$\varepsilon = \sqrt{1+4\omega}.$$

---

[15] Since (129) only applies for $t = s$ the value of y must be computed by recurrence from $t = -1$ to $t = s - 1$.

Developing $(1+\varepsilon)$ and $(1-\varepsilon)$ by the binomial formula the bracket in (131) becomes

$$\sum_{i=0}^{t+1}\binom{t+1}{i}(\varepsilon^i-(-\varepsilon)^i)=\sum_{i=0}^{t+1}\binom{t+1}{i}\varepsilon^i(1-(-1)^i) \quad \ldots \quad (132)$$

which reduces to
$$\eta_t=\left(\frac{\mu}{2}\right)^t\sum_{i=1,3,5\,\ldots}\binom{t+1}{i}(1+4\omega)^{(i-1)/2}. \quad \ldots \quad (133)$$

This expression is identically equal to the coefficient of Y_0 in (35). For $t=3$ we get for instance by (133)

$$\left(\frac{\mu}{2}\right)^3\left[\binom{4}{1}+\binom{4}{3}(1+4\omega)\right]=\mu^3(1+2\omega) \quad \ldots \quad (134)$$

which checks with the coefficient of Y_0 on the third line in (8). Similarly the coefficient of Y_{-1} in (35) is what is obtained by replacing t by $(t-1)$ in (133) and multiplying by ν. (cf. the second term in the right member of (130)).

A warning note in connection with the general solution of the complete Hicksian difference equation. One sometimes encounters analytical work where the reasoning runs as if the characteristic roots were sufficient to describe the time shape of the solution of a complete, i.e. non-homogeneous, difference equation. This is not so. The characteristic roots described the time shape of the *cumulator,* not that of the time-function Y_t which is a solution of the complete, i.e. non-truncated, difference equation. The solution Y_t is indeed *not yet determined* if only the roots are known. The time shape of Y_t may be anything depending on the time shape of the additional term H_t in the difference equation (5). This is only another aspect of the fact that (5) is one equation in two variables Y and H, hence one degree of freedom.

Case II. *A real double root of the characteristic equation corresponding to the truncated Hicksian difference equation.* In this case the argument in (131) does not hold good because $(\lambda_1-\lambda_2)$ vanishes. But the expression (133) can be used also in this case because the denominator $(\lambda_1-\lambda_2)$ which is equal to $\mu\sqrt{1+4\omega}$, has been *removed* when we pass from (131) to (133). Compare also (149).

The case $\lambda_1=\lambda_2$ is obtained from (133) simply by putting $1+4\omega=0$, (cf. (103)-(104)). Only one term will then be left in (133) namely the term corresponding to $i=1$. This gives

$$\eta_t=(t+1)\left(\frac{\mu}{2}\right)^t \quad \text{(when } \lambda_1=\lambda_2\text{)}. \quad \ldots \quad (135)$$

This expression inserted into (129) will give the general solution of the complete difference equation in the case of two equal roots (the roots are necessarily real when they are equal).

The same expression (135) for the cumulator is obtained by following the general rule (56). This leads to

$$\eta_t=c_1\left(\frac{\mu}{2}\right)^t+c_2t\left(\frac{\mu}{2}\right)^t. \quad \ldots \quad (136)$$

Determining here the constants c_1 and c_2 in such a way that (136) satisfies the initial conditions (125), we get back to (135).

Degeneration to a first order difference equation. The general solution in the case where the difference equation degenerates into a first order equation because of

$$\nu = 0 \qquad \dots \quad (137)$$

follows immediately from (35) by putting $\omega = 0$. This leaves in the first and third brackets of (35) only one term, namely, the one for $i = 0$, while the second bracket vanishes. This gives

$$Y_t = \mu^t Y_0 + \sum_{\tau=1}^{t} \mu^{t-\tau} H_\tau \qquad \dots \quad (138)$$

which can also be written

$$Y_t = \mu^t Y_0 + \sum_{\tau=0}^{t-1} \mu^\tau H_{t-\tau} \quad (t = 1, 2, \dots, \infty, \text{ when } \nu = 0). \qquad \dots \quad (139)$$

The latter expression corresponds to (129) for $\nu = 0$, $s = 1$ and $\eta_t = \mu^t$.

The above simple remarks which immediately lead to (135) and (139) show the usefulness of having different expressions for the solution. The whole situation is under more complete control when we approach the problem both by recurrence and by means of characteristic roots.

THE PROGRAMMING APPROACH

Smoothness and rapidity. The programming approach within a given time horizon $t = 1, 2, \dots, T$ may be formulated in different ways.

A common feature of all these formulations is that we want to produce a development of national income Y_t which is in some sense "smooth" and at the same time "rapid". In practice other desiderata may also be implied but for the time being let us assume that the "smoothness" and "rapidity" of the development Y_t are our desiderata.

The "smoothness" of Y_t will depend on two things : the nature of the cumulator η_t as defined by (131) and (125) and on the possible vicissitudes of H_t over time. The way in which "smoothness" depends on the nature of the cumulator can be disposed of simply by saying that the cumulator η_t should not be wave shaped.[16] This is obtained simply by imposing (114), which is equivalent to (115). This is only a single *bound* on the government strategy parameters. In some of the programming formulations the effect of vicissitudes in H_t may be taken account of in special ways exemplified in the sequel.

The "rapidity" of the development of a time-function offers no definitional difficulty in the case where the function is simply an exponential in t, because in this case its relative rate of growth is constant. But in our problem the relative rate will as a rule not be constant, (cf. (127) and 129)). We therefore face the question whether we want to attribute most importance to a rapid growth in the near future or most importance to the final achievements in the long run. This is a matter to be decided by the policy maker (whether it is a democratic parliament or some dictator). It does not restrict generality if we let $t = 0$ denote the *decision time point* (the plan making point of time if a true planning procedure is envisaged, otherwise it would simply be a point of time when discussion takes place).

---

[16] Cf. the remarks after (148) as well as (155).

The policy maker's decision on the relative preference between the present and the future may be expressed by a preference function of the form

$$F = \sum_{t=1} P_t Y_t = \max. \qquad \dots \quad (140)$$

where the P_t are the policy maker's preference coefficients, and T a time horizon adopted for the repercussions to be taken account of. T is the *repercussion period.* To put $T = \infty$ is sheer abstractism based on the assumption of constant structural coefficients and strategy coefficients over an indefinite future. In practice a much shorter repercussion period will have to be considered.

A particularly simple form of (140) would be to put all P_t equal to zero except P_T. In this case the preference function is simply

$$F = P_T Y_T. \qquad \dots \quad (141)$$

Another particularly simple form of (140) is

$$P_t = d^t \qquad \dots \quad (142)$$

where d (generally positive and less than unity) is a discount factor. This discount factor would (in an economy which is not completely ruled by the time honoured unenlightened financialism) have nothing to do with any "market rate of interest" but would simply be an expression for the policy maker's desires.

More complicated forms of the preference function than the linear form (140) may be considered for special purposes.

So much for the desiderata to be achieved. Differences in the choice of the *means to be used* (and to be analyzed in terms of the model) will result in different programming formulations.

Formulation 1 : Let the initial conditions Y_0, Y_{-1} as well as the parameters μ and ω be *given* (the latter being given through known constant values of the structural parameter α, β and through such decided upon values of the strategy coefficients κ, λ as will satisfy (114)).

In such a situation at time 0 how can the autonomous investments H_1, H_2 $\dots$, H_T be chosen so as to obtain a "smooth" and at the same time "rapid" development of Y_t within the horizon T ?

In this formulation the application of (35) or, if we like (129) must be forward looking, i.e. the point of time 0 must be today and $t = 1, 2, \dots$ must be points of time in the future.

This formulation of the programming problem has an obvious interest in the case where government is at least to some degree able to *influence* the course of H

in the future. And even in the case where government has left autonomous investment largely in the hands of private business, it might be interesting for government to know what the nationally desirable evolution of autonomous investment is, and to measure how far the statistically observed action of private business is from achieving the nationally desirable optimum. So in any case the present formulation of the programming problem might be interesting.

If the preference function is of the linear form (140) we see, by inserting from (129), that F will also be linear in the T variables $H_1, H_2 \dots H_T$. Hence so far as the preference function is concerned the programming problem is linear.

So far as the bounds are concerned we would in a realistic analysis have to impose the T conditions.

$$\underline{C}_t \leqslant C_t \quad (t = 1, 2 \dots, T) \qquad \dots \quad (143)$$

where $\underline{C}_t$ $(t = 1, 2 \dots, T)$ are lower bounds for consumption.

These lower consumption bounds may be taken as decisionally *given*, i.e. as magnitudes not to be determined through the programming analysis. If so, the policy maker would have to decide on how stringent these bounds should be (perhaps near to the physical minimum of existence in the beginning of the repercussion period). If the bounds $\underline{C}_t$ are given, and if we insert from (114) and make use of the known initial condition Y_0, the bounds (143) will also emerge as linear in the decisional variables $H_1, H_2 \dots, H_T$.

If the lower bounds $\underline{C}_t$ are not taken as given in an absolute sense but determined as certain given *fractions* of Y_t, the bounds (143) would again become linear in terms of the decisional variables $H_1, H_2 \dots, H_T$. The same is even true if each $\underline{C}_t$ is determined as a linear function of all the $H_1, H_2 \dots, H_T$ with given constants.

In all these cases the bounds (143) will be linear.

But (143) are not the only bounds that must realistically be taken account of. One must *at least* consider the bounds springing from the need to consider *production capacities*, (cf. my introductory remarks on the realism of the model).

One simple way in which to take account of this type of bounds would be to use the Harrod-Domar idea in the following way.

Let $\overline{Y}_t$ be the production capacity as determined by the size of real capital in the point of time t. This means that we must have

$$Y_t \leqslant \overline{Y}_t \qquad (t = 1, 2 \dots, \infty). \qquad \dots \quad (144)$$

By Harrod-Domar reasoning $\overline{Y}_t$ would be of the form

$$\overline{Y}_t = \sigma K_t \qquad (t = 1, 2 \dots, \infty) \qquad \dots \quad (145)$$

where K_t is real capital at time t, and σ is the *output-to-capital ratio*, supposed structurally given and constant over time (which means that we refrain from considering the specific technological effects of new investments on σ).

Since by definition

$$K_t = K_0 + \sum_{\tau=1}^{t} (I_\tau + H_\tau) \qquad (t = 1, 2 \ldots, \infty) \qquad \ldots \quad (146)$$

where K_0 is the initial size of real capital, we get from (144)-(145)

$$Y_t \leqslant \sigma \Big[K_0 + \sum_{\tau=1}^{t} (I_\tau + H) \Big]. \qquad \ldots \quad (147)$$

This again is, by (2) and (129) a bound which is linear in the decisional parameters $H_1, H_2 \ldots, H_T$.

The introduction of the two new variables $\bar{Y}_t$ and K_t does not change the number of degrees of freedom since we also have two new equations, namely (145) and (146).

If we want to introduce labour requirement we may do it in the form of the bound

$$Y_t \leqslant \phi \bar{N}_t \quad (t = 1, 2 \ldots, \infty) \qquad \ldots \quad (148)$$

where ϕ is a structurally given labour coefficient and $\bar{N}_t$ a given forecast of the working population (so far without taking account of the problem of specialized skilled labour, which is so conspicuously important, not least in developing countries). In this case too the bound (148) is linear. And the inclusion of unemployment $(\bar{N}_t - Y_t/\phi)$ as part of the preference function would not change the linear character of this function.

So, in all we get a programming model which is completely linear both with regard to preference function and bounds. It may therefore be solved by one of the known methods.[17]

If we want to consider a smooth course of the H_t as an additional desideratum besides the main desideratum expressed in (140) we may add for instance $\Big[-P_x \sum_{t=1}^{T} (H_t - H_x)^2 \Big]$ to (140), where P_x is a positive preference coefficient and $H_x = \frac{1}{T} \sum_{t=1}^{T} H_t$ the average of H_t over the repercussion period. The preference function F thus obtained is concave (since the matrix of its second order derivatives will be negative definite). The bounds remain linear. The ensuing programming problem is therefore not a particularly difficult one.[18] Any other additional term which retains the concave character of the preference function could be considered.

---

[17] Since we will in this case have a moderate number of degrees of freedom, namely T, the number of decisional variables $H_1, H_2 \ldots, H_T$, but a much larger number of dependent variables, the problem is one where my multiplex method will work to great advantage. It is coded by Mr. Ole-Johan Dahl for use on the electronic computer of the Norwegian Defence organization, and has been very successfully used here for rather large size problems.

[18] For problems with linear (upper and/or lower) bounds and a concave preference function the multiplex method works very well, as I reported on at the 1960 Tokyo meeting of the International Statistical Institute. The paper is printed in the report of the meeting.

Formulation 2 : This is similar to Formulation 1 but $H_1, H_2 \ldots, H_T$ are now not deterministically defined but defined through the simultaneous probability distribution of $H_1, H_2 \ldots, H_T$. The technical aspect of this programming solution in this case will in principle be somewhat similar to that under Formulation 1, except for the fact that the decisional parameters will now be the parameters that determine the simultaneous distribution of $H_1, H_2 \ldots, H_T$. It goes, however, without saying that this stochastic twist of the problem will in practice increase the computational difficulties very much.

Formulation 3 : This formulation is similar to the Formulation 1, in so far as the magnitudes $H_1, H_2 \ldots, H_T$ are also now considered as deterministic parameters to be determined by the programming procedure but in addition the government strategy parameters κ and λ are also considered as decisional, i.e. as variables to be decided upon through the programming analysis, subject however to the condition (114).

Since the parameters κ and λ—through μ and ω—enter in an extremely non-linear manner in (35) and (129), and hence by (1) in the corresponding expression for consumption $C_1, C_2 \ldots, C_T$, real capital $K_1, K_2 \ldots, K_T$ etc., the computational difficulties will now be great, but possibly not insurmountable for a moderately large horizon T. (cf. the remarks on iteration at the end of this paper).

Formulation 4 : Same as Formulation 3 except for the fact that $H_1, H_2 \ldots, H_T$, are now considered not as deterministically defined but as stochastically defined through the simultaneous distribution of $H_1, H_2 \ldots, H_T$.

This problem will be computationally very hard to handle. It is very doubtful whether an attempt in this direction is worth the trouble, particularly in view of the fact that it will probably not lead to any essentially new and practically more useful results than those that could be obtained through a solution of the Formulation 3.

Formulation 5 : *Solution when autonomous investment is not under government control.* If autonomous investment is not under government control, the only thing government can do, within the framework of the model (1)-(4), is to handle the strategy parameters (κ, λ) in some way which is as "optimal" as it can be made without knowing the autonomous investment.

A plausible thing to do under these circumstances is to consider a (κ, λ) strategy which is optimal on the assumption that there will be *no* autonomous investments in the future i.e. for $t = 1, 2 \ldots \infty$ ($t = 0$ being "today" i.e. the day when the decision about (κ, λ) is made). Any actually occurring autonomous investment will then simply come as an *addition* to the national income which is theoretically produced under the strategy considered.[19]

We denote the solution of the difference equation (5) under this assumption by the letter y_t. By (131) we have

$$y_t = \eta_t Y_0 + \nu\eta_{t-1} Y_{-1} \quad (t = 1, 2 \ldots, \infty) \qquad \ldots \quad (149)$$

where Y_0 and Y_{-1} are the *actually realized* magnitudes of the national income in years 0 and (-1) respectively regardless how those magnitudes have been achieved.

---

[19] As an alternative we may assume that the H_t are given through some sort of *estimate* made *independently* of the model. This would make the computations more complicated but would not cause insurmountable difficulties.

As before, we assume provisionally the non-wave case for the cumulator. This is equivalent to imposing the condition (114) which is the same as (115). It will turn out that this is not a restrictive assumption since the non-wave case is actually the optimal solution.

If the two characteristic roots (provisionally real by assumption) are different, the expression for the cumulator is (127). This expression can be transformed in such a way that it becomes applicable even in the case $\lambda_1 = \lambda_2$. We simply divide by $(\lambda_1 - \lambda_2)$ which gives

$$\eta_t = \sum_{v=0}^{t} \lambda_1^{v}\lambda_2^{t-v} = \sum_{v=0}^{t} \lambda_1^{t-v}\lambda_2^{v} \qquad (\lambda_1 = \lambda_2 \text{ or } \lambda_1 \neq \lambda_2). \qquad \ldots \quad (150)$$

The two expressions in (150) for η_t are equivalent.

Since by (118) the expression (149) can be written

$$y_t = \eta_t Y_0 - \lambda_1\lambda_2\eta_{t-1}Y_{-1} \quad (t = 1, 2 \ldots, \infty) \qquad \ldots \quad (151)$$

we can by (150) write the expression for y_t in either of the two following forms

$$y_t = Y_0 \sum_{v=0}^{t} \lambda_1^{v}\lambda_2^{t-v} - Y_{-1} \sum_{v=0}^{t-1} \lambda_1^{v+1}\lambda_2^{t-v} \quad (\lambda_1 = \lambda_2 \text{ or } \lambda_1 \neq \lambda_2) \quad \ldots \quad (152)$$

$$y_t = Y_0 \sum_{v=0}^{t} \lambda_1^{t-v}\lambda_2^{v} - Y_{-1} \sum_{v=0}^{t-1} \lambda_1^{t-v}\lambda_2^{v+1} \quad (\lambda_1 = \lambda_2 \text{ or } \lambda_1 \neq \lambda_2) \quad \ldots \quad (153)$$

We may, if we like, consider the programming formulation as one in (λ_1, λ_2). Indeed, to any set (λ_1, λ_2) corresponds by (117)-(118) one and only one set (μ, ν), and inversely to any set (μ, ν) corresponds a set (λ_1, λ_2) which is uniquely determined, apart from an interchange of the numbering of the two roots λ_1 and λ_2 (we may, if we like, conventionally choose the numbering so that $\lambda_1 \geqslant \lambda_2$).

Considering the programming problem as a problem in (λ_1, λ_2) we take the partial derivatives of y_t with respect to λ_1 and λ_2 respectively. When determining the former of these two derivatives it is most convenient to use (152) and when determining the latter derivative it is most convenient to use (153).

After some transformations we get

$$\frac{\partial y_t}{\partial \lambda_1} = \left[\sum_{v=1}^{t} v\lambda_1^{v-1}\lambda_2^{t-v}\right](Y_0 - \lambda_2 Y_{-1}) \qquad \ldots \quad (154)$$

$$\frac{\partial y_t}{\partial \lambda_2} = \left[\sum_{v=1}^{t} v\lambda_1^{t-v}\lambda_2^{v-1}\right](Y_0 - \lambda_1 Y_{-1}). \qquad \ldots \quad (155)$$

This shows that we can determine values of λ_1 and λ_2 that are independent of t and such that they *annihilate* both derivatives simultaneously. These values are simply obtained by putting the last parenthesis in (154) and (155) equal to zero. This gives

$$\lambda_1 = \lambda_2 = \frac{Y^0}{Y_{-1}}. \qquad \ldots \quad (156)$$

Without going into a detailed algebraic proof we see from numerical examples, (cf. tables (180) and (181)), that (156) actually yields a maximum.

The single parameter strategy. We are thus led to the case of two equal roots. Since the two roots cannot be equal unless they are real and since the above argument is valid *regardless* of whether λ_1 and λ_2 are looked upon as real or complex, we have actually proved that the programming in λ_1 and λ_2 leads to the case of two equal and real roots. That is to say we must have the Case II in (106).

The programming problem *is thus reduced to a single-parameter problem.*

Let
$$r = \frac{\mu}{2} \quad \text{(the general case of two equal roots)} \qquad \dots \quad (157)$$

be the common value of the two roots. The parameter r in (157) has the same meaning as in (64) and (69). The problem is to determine the parameter r in such a way that y_t becomes as large as possible when Y_0 and Y_{-1} are given. The value of r which achieves this, follows directly from (156). We denote this value by the subscript 1. That is, we have

$$r_1 = \frac{Y_0}{Y_{-1}} \text{(the optimal value of the two equal roots).} \qquad \dots \quad (158)$$

The symbol r_1 can be looked upon as designating both the optimum value of the strategy parameter r and the initially observed *actual growth factor.*[20]

The cumulator in the case of two equal roots is easily derived from (135) and (157). We get

$$\eta_t = (t+1)r^t \quad (t = 1, 2, \dots, \infty) \quad (r \text{ arbitrary, not necessarily equal to } r_1). \quad \dots \quad (159)$$

This expression can also be obtained by the general method of (56), (89)-(91).

From (117)-(118) and (157) we now get

$$\mu = 2r \qquad \dots \quad (160)$$

$$\nu = -r^2. \qquad \dots \quad (161)$$

So that by (6)-(7)

$$\kappa = -\alpha + 2r - r^2 \quad \text{(for any } r\text{)} \qquad \dots \quad (162)$$

$$\lambda = \beta - r^2 \quad \text{(for any } r\text{)} \qquad \dots \quad (163)$$

The expression (162) shows how much larger the strategy parameter κ must be chosen than the structural constant α in the Case II in (108) when $(\beta - \lambda)$ has been chosen.

Inserting (158) into (162)-(163) we get the optimal strategy values

$$\kappa_0 = -\alpha + 2r_1 - r_1^2 \quad \text{(optimal)} \qquad \dots \quad (164)$$

$$\lambda_0 = \beta - r_1^2 \quad \text{(optimal)} \qquad \dots \quad (165)$$

where r_1 is given by (158).

---

[20] When I speak here of the "growth factor" (which is unity plus the growth rate) I take the word factor simply *in* the sense of "a number by which to multiply", not in the sense of "something which can explain." In the latter sense the word factor is used *for instance in biology* when one speaks of a "growth factor" or "factor for growth."

In any *realistic case in economics* r_1 determined by (158) will be *positive*, so that we do not get the zig-zag case (the case where η_t changes from plus to minus in every other point of time).

Government current account expenditure is given by (3) which we can also write

$$G_t = (\kappa-\lambda)Y_{t-1}+\lambda Y_{t-2} \quad (t = 1, 2, \ldots, \infty). \qquad \ldots \quad (166)$$

Inserting (162)-(163) in (166) we get

$$G_t = [2r-(\alpha+\beta)]Y_{t-1}+[\beta-r^2]Y_{t-2} \qquad \ldots \quad (167)$$

(for any r in the case of two equal roots)

If we insert the optimal strategy value (158) for r in (167) we get the *optimal* government current account expenditure in year t, when Y_{t-1} and Y_{t-2} are the *actual* national incomes in year $(t-1)$ and $(t-2)$ regardless of the way in which these magnitudes happen to have emerged through previous autonomous investments. This is the meaning of our determining (by a decision at time $t = 0$) the optimal strategy parameters κ_0 and λ_0 *as if* all the H_1, H_2 ..., H_{t-1} are going to be zero.

In the general case of a double root—i.e. with an arbitrary value of r we have

$$\eta_{t+1}-\eta_t = r^t[r(t+2)-(t+1)] = \eta_t\left[r\,\frac{t+2}{t+1}-1\right] \qquad \ldots \quad (168)$$

(for any r in the case of two equal roots)

and by (129) $$Y_{t+1}-Y_t = H_{t+1}+\Big[r^t[r(t+2)-(t+1)]Y_0-r^{t+1}[r(t+1)-t]Y_{-1}\Big]$$

$$+\sum_{\tau=0}^{t-s} r^\tau\Big[r(\tau+2)-(\tau+1)\Big]H_{t-\tau} \qquad \ldots \quad (169)$$

$$(t = s, s+1, \ldots, \infty); \quad (s = 1, 2, \ldots, \infty).$$

The middle term in (169), i.e. the big bracket represents *the after-effect of the initial conditions*, the last term, i.e. the summation over τ represents the *after-effect of the autonomous investments* up to t inclusive and the first term in (169) represents *the effect of the new autonomous investments* at time $t+1$.

As t tends toward infinity we get from (168)

$$\frac{\eta_{t+1}-\eta_t}{\eta_t} = (r-1) \qquad \text{(as } t\to\infty \text{ for any value of } r\text{)}. \qquad \ldots \quad (170)$$

By the passage to $t\to\infty$ and using the formulae (129), (168), (170) and (159) become

$$\frac{Y_{t+1}-Y_t}{Y_t} = (r-1)+\frac{H_{t+1}}{Y_t}+\frac{r\sum_{\tau=0}^{t-s} r^\tau H_{t-\tau}}{Y_t} \quad \text{(as } t\to\infty \text{ for any value of } r\text{)} \qquad \ldots \quad (171)$$

(170) and (171) show that *asymptotically* the rate of growth of the cumulator will be $(r-1)$—as it would be for *any* t if the cumulator were a simple exponential—, and

asymptotically the rate of growth of national income itself will be $(r-1)$ plus two correctional terms. The first is due to H_{t+1}, which we can in principle choose arbitrarily inside the model (1)-(4). And the second correctional term represents the after-effect of previous autonomous investments. All we can say about the latter of the correctional terms is that they will be positive if the autonomous investment are on the average positive. Even if $H_{t+1} = 0$ it will therefore be safe to assume that asymptotically the growth rate of Y_t will be *at least* $(r-1)$.

More important than to study the passage $t \to \infty$ is to see what happens in the first few years after the point of time 0. I will illustrate this below by some numerical examples. They will also show the effect of changing r. These examples will verify the above theoretical results. We consider the case where all the autonomous investment are zero.

For any value of r (i.e. assuming two equal roots) we have by (160)-(161)

$$y_t = 2ry_{t-1} - r^2 y_{t-2} \quad \dots \quad (172)$$

(the recurrent formula in the case of two equal roots),

$$y_t = (t+1)r^t Y_0 - tr^{t+1} Y_{-1} \quad \dots \quad (173)$$

(the ab-initio formula in the case of two equal roots).

By (167) the government expenditure on current account will, on the assumption of no autonomous investments, be

$$g_t = [2r-(\alpha+\beta)]y_{t-1} + [\beta - r^2]y_{t-2} \quad (t = 1, 2, \dots, \infty \text{ for any } r). \quad \dots \quad (174)$$

For $t = 1$ we will in (174) have to put y_0 and y_{-1} equal to respectively Y_0 and Y_{-1} where Y_0 and Y_{-1} are the magnitudes actually realized in these two points of time.

From (173) we get

$$\frac{dy_t}{dr} = t(t+1)r^{t-1}(Y_0 - rY_{-1}) \quad \dots \quad (175)$$

and using this in (174) we get

$$\frac{dg_t}{dr} = \Big[[2r-(\alpha+\beta)]r(t-1)t + [\beta - r^2](t-2)(t-1) + 2r^2 \Big] r^{t-3}(Y_0 - rY_{-1}). \quad \dots \quad (176)$$

Comparing (175) and (176) we see that during the variation of r, g_t and y_t must have extremum for the same value of r namely the value r_1 defined by (158).

By inserting the ab-initio formulae (173) for y_{t-1} and y_{t-2} we can derive the corresponding ab-initio formula for g_t, but this is unnecessary for our purpose.

Numerical examples. In our numerical examples we use

$$\alpha = 0.8 \qquad \beta = 2. \quad \dots \quad (177)$$

And we consider two sets of initial conditions, namely

Example A : $Y_{-1} = 100$ and $Y_0 = 103$ ... (178)
Example B : $Y_{-1} = 100$ and $Y_0 = 107$. ... (179)

The result for different values of r are given in (180) and (181) respectively.

Example A: $Y_{-1} = 100$ and $Y_0 = 103$

	$t=1$			$t=2$			$t=3$		
alternative r	y_t	g_t	$\frac{100g_t}{y_t}$	y_t	g_t	$\frac{100g_t}{y_t}$	y_t	g_t	$\frac{100g_t}{y_t}$
1.00	106.00	17.60	16.60	109.00	18.20	16.70	112.00	18.80	16.79
1.01	106.05	17.65	16.64	109.15	18.21	16.68	112.30	18.78	16.72
1.02	106.08	17.68	16.66	109.24	18.22	16.67	112.49	18.77	16.69
1.03	106.09	17.69	16.67	109.27	18.22	16.67	112.55	18.77	16.67
1.04	106.08	17.68	16.66	109.24	18.21	16.67	112.49	18.77	16.69
1.05	106.05	17.65	16.64	109.15	18.21	16.68	112.29	18.77	16.72
1.06	106.00	17.60	16.60	108.99	18.19	16.68	111.96	18.79	16.78
1.07	105.93	17.53	16.55	108.77	18.16	16.69	111.48	18.80	16.86
1.08	105.84	17.44	16.48	108.48	18.12	16.70	110.85	18.80	16.96
1.09	105.73	17.33	16.39	108.12	18.07	16.71	110.08	18.81	17.09
1.10	105.60	17.20	16.29	107.69	18.01	16.72	109.14	18.81	17.23
1.20	103.20	14.80	14.34	99.36	16.40	16.51	89.86	18.05	20.09
1.30	98.80	10.40	10.53	82.81	12.18	14.71	48.33	14.07	29.10
1.40	92.40	4.00	4.33	56.84	4.12	7.25	-21.95	3.70	-16.84
1.50	84.00	-4.40	-5.24	20.25	-8.95	-44.20	-128.25	-16.95	13.22

... (180)

Example B : $Y_{-1} = 100$ and $Y_0 = 107$

1.00	114.00	14.40	12.63	121.00	15.80	13.06	128.00	17.20	13.44
1.01	114.13	14.53	12.73	121.39	15.83	13.04	128.79	17.15	13.32
1.02	114.24	14.64	12.82	121.73	15.85	13.02	129.47	17.11	13.22
1.03	114.33	14.73	12.88	122.00	15.88	13.02	130.03	17.08	13.14
1.04	114.40	14.80	12.94	122.22	15.90	13.01	130.48	17.07	13.08
1.05	114.45	14.85	12.98	122.38	15.92	13.01	130.81	17.05	13.04
1.06	114.48	14.88	13.00	122.47	15.93	13.01	131.01	17.05	13.01
1.07	114.49	14.89	13.01	122.50	15.93	13.01	131.08	17.05	13.01
1.08	114.48	14.88	13.00	122.47	15.93	13.01	131.01	17.05	13.01
1.09	114.45	14.85	12.98	122.37	15.91	13.00	130.80	17.05	13.04
1.10	114.40	14.80	12.94	122.21	15.89	13.00	130.44	17.05	13.07
1.20	112.80	13.20	11.70	116.04	14.80	12.69	117.50	16.51	14.05
1.30	109.20	9.60	8.79	103.09	11.33	10.99	83.49	13.23	15.85
1.40	103.60	4.00	3.86	80.36	4.28	5.33	21.95	4.14	18.88
1.50	96.00	-3.60	-3.75	47.25	-7.55	-15.98	-74.25	-14.55	19.60

... (181)

These tables show clearly that $r_1 = \frac{Y_0}{Y_{-1}} = 1.03$ is the optimum strategy in the first example, and $r_1 = \frac{Y_0}{Y_{-1}} = 1.07$ the optimum strategy in the second example.

The optimum relative government expenditure on current account. In the case of the optimal strategy we get

$$y_{t.0} = r_1^t Y_0 \qquad \text{... (182)}$$

$$g_{t.0} = r_1^{t-2} Y_0[r_1^2-(\alpha+\beta)r_1+\beta] \qquad \text{... (183)}$$

where the subscript 0 on $y_{t.0}$ and $g_{t.0}$ indicates the result of the optimum choice (158).

From (182)-(183) we get immediately

$$\frac{g_{t.0}}{y_{t.0}} = 1 - \frac{\alpha+\beta}{r_1} + \frac{\beta}{r_1^2} \qquad \ldots \quad (184)$$

and hence

$$\frac{d\left(\dfrac{g_{t.0}}{y_{t.0}}\right)}{dr_1} = \frac{\alpha+\beta}{r_1^3}\left(r_1 - \frac{2\beta}{\alpha+\beta}\right). \qquad \ldots \quad (185)$$

The formula (184) shows that the optimum *relative* government expenditure—under the conditions we have assumed for the optimum—*is independent of t*. But it will change with the initial conditions (which determine r_1), and the way this change takes place is given in (185).

It is an interesting fact that the above optimalization process is *independent of t*. This means that in *any* point of time $(t-1)$ we can determine the optimum value of the two strategy parameters $(\kappa-\lambda)$ and λ, as if it were only a question of achieving the best result in the *next year t*, assuming that no autonomous investment is to take place in the next year. Whatever autonomous investment that is *actually going* to be achieved in the next year t will only produce so much additional increase in the national income Y_t. The development will therefore *gain momentum* from year to year by any autonomous investments which the nation's real resources will permit and the private investors choose to make. The actual growth process will therefore be more rapid than that illustrated in (180) and (181). Each year will so to speak represent a change of initial conditions, and hence a new optimal value for the parameter r.

GOVERNMENT CURRENT ACCOUNT BUDGET STRATEGY UNDER NON-CONTROLLED INVESTMENTS

The above consideration leads to the following rule for a *moving optimalization*.

Moving optimalization. Each year t the *actually realized* national income in year $t-1$ and $t-2$ will be known. From these data Y_{t-1} and Y_{t-2} one should compute the optimum factor r_t, determined by

$$r_t = \frac{Y_{t-1}}{Y_{t-2}} \quad \text{(The optimum factor computed at } t\text{)} \quad (t = 1, 2, \ldots, \infty). \qquad \ldots \quad (186)$$

This value of r_t is inserted for r in (167). This gives the optimum government current account expenditure in year t.

That is, in year t the government current account expenditure ought to be put equal to the optimum value

$$G_{t.0} = [2r_t-(\alpha+\beta)]Y_{t-1}+[\beta-r_t^2]Y_{t-2} = Y_{t-1}[r_t+\frac{\beta}{r_t}-(\alpha+\beta)] \qquad \ldots \quad (187)$$

where r_t is given by (186).

This figure will, of course, in practice only indicate a *guiding signal* for government expenditures. It has been computed only by considerations on the general growth rate of the economy. Other considerations (social justice, defence etc.) may motivate an action whereby one sacrifices part of the overall national growth rate.

But this being said, I believe that in the absence of a real control over the autonomous investments in the nation, and of a complete programming solution involving also the magnitudes H_1, H_2 ..., the above expenditure rule is as good a rule as can reasonably be deviced. But the rule only leads to protection against decline, not to real *progress* as will be seen from the subsequent argument.

By the right hand expression in (187) the optimal expenditure ratio $G_{t.0}/Y_{t-1}$ *depends only on the recent growth rate* r_t (and on the two structural coefficients α and β). Furthermore, the way in which the optimal expenditure ratio depends on r_t transpires immediately from (187) and from the expression for the derivative

$$\frac{d\left(\frac{G_{t.0}}{Y_{t-1}}\right)}{dr_t} = 1-\frac{\beta}{r_t^2} = \text{negative in realistic cases.} \quad \dots \quad (188)$$

Since we may assume α and β positive—and in realistic case α smaller than unity and β larger than unity—the optimal expenditure ratio is by (187) very large for very small positive values of r_t. With increasing r_t it decreases, passing the value $(1-\alpha)$ for $r_t = 1$, continuing to decrease and reaches its minimum value $(1-\alpha)-(\sqrt{\beta}-1)^2$ for $r_t = \sqrt{\beta}$. From this point on it increases without interruptions as r_t continues to increase.

The most realistic range for r_t is between 1 and $\sqrt{\beta}$ and in this range we can draw the following conclusion : The lower the growth rate for national income has been in recent years the *larger* should the optimum government expenditure on current account be in relation to national income. This conclusion is rather at variance with the favourite conservative argument about the advantage of low government expenditure. This need to *increase* government expenditure (in order to obtain optimality) is all the smaller the larger β, i.e. (188) *more* negative when β is large.

Protection against decline, not a driving force for progress. If no autonomous investments accrue, the growth rate of national income (under optimal government current account expenditure at all times) is *constant* and determined by some distant initial conditions (cf. 182). In the model (1)-(4) the optimalization of government current account expenditure is therefore only a *protection against decline.* It protects what we had, namely r_1 (or r_t in moving optimalization), but is no driving force for progress. Only through autonomous investments can the growth rate be increased. And when it is increased, the optimal expenditure ratio should (in order to protect the new growth rate) be adapted in the way we have just discussed in connection with (187) and (188). If the model should have been able to display a driving force for progress, it would have been necessary to add a relation that indicates an effect of government current expenditure on autonomous investment.

Attempts at reaching higher growth rates by influencing autonomous investment. If account is taken of bounds such as (143), (144) and (148), the more complete programming Formulations 1-4 may be realistic enough to make it possible to reach much higher growth rates. But then autonomous investments must be controlled.

One may simplify the computation by an iterative procedure which consists in first determining the optimal strategy parameter r_t by (186) and then proceeding to the solution of the Formulation 1 problem, which is simply a linear programming problem. Having solved this problem one may reshape Formulation 5 so as to take account of the autonomous investments now determined. Thus we reach an improved value for r_t. This improved value may again be taken as the basis for a programming problem of type 1. And so on. This might lead to an approximate solution to the Formulation 3 problem and would give *coordination* of current account and investment policies.

The need for a more complete planning model. But this whole approach is as yet too simple to be taken as a complete planning model, because it does not treat explicitly such important aspects as *import-export* effects (balance of payments), *sectorial breakdown, substitution possibilities* among sector inputs (i.e. not all the input coefficients being constants), the *infra effect* of investments in changing the coefficients of a model, the problem of *education* and *skilled labour* (which formally can be handled very much in the same way as the problem of capital formation and with bound-effects similar to that produced by material capital). All these problems are basically important for planning in an underdeveloped country. Nor can they be neglected in a more advanced country which will introduce a certain amount of planning in its economy.

In this connection I shall not discuss the various attempts I have made of developing a complete planning model.[21]

SUMMARY

An explicit general solution of the Hicksian Model, holding true for any period of time t and with any arbitrarily given right member (the term which makes the difference equation non-homogeneous) is given. The solution has one degree of freedom (one arbitrary function of time) when the parameters of the model are given.

Programming techniques are indicated either to determine desirable values of the parameters in the Hicksian Model or to choose one of the variables as a decisional function of time. In both cases, we may haze a linear or a more general preference function having the aim of minimising fluctuations in incomes from their long run growth path and/or assuring rapid growth in the long run.

There are five formulations, making different assumptions about the time shape of autonomous investment. One formulation applies to the case where the autonomous investment is not under government control.

The recurrence method as well as the method of characteristic roots for solving a linear difference equation of any order with constant co-efficients and an arbitrary right member (the non-homogeneous case) are stated in a precise and easily understandable form with special application to the Hicksian Model.

---

[21] See for instance, "An implementation system for optimal national economic planning without detailed quantity fixation from a central authority". Memorandum, 3 January 1963 from the Oslo University Institute of Economics. Soon to appear in print.

THE NATIONAL BALANCE SHEET OF THE SOVIET UNION[1]

By RAYMOND W. GOLDSMITH

Yale University, U.S.A.

I

Discussions or investigations of the monetary or financial aspects of the Soviet economy are extremely rare with the exception of certain aspects of the Soviet tax system that affect the pricing process. Moreover, the attention devoted to the financial aspects of the Soviet economy seems to have declined rather than increased in the course of time. This state of affairs is astonishing, although it may be explained to some extent by the general lack of serious comparative studies of financial structure and development. One should think that the principles and techniques applied to the financial side of its economy by the country which first developed a centrally planned economy with ownership by the state of virtually all means of production and which has operated this system longer than any other country, would be of great interest both to economists who try to understand the Soviet economic system, and even more to those who attempt to learn from the comparison of the ways in which the economies of different countries operate. For the Soviet Union has always, apart possibly from a short period of war communism, possessed a reasonably comprehensive financial system with most of the paraphernalia usually associated with it: currency and deposit money; central, short term credit, long term credit and savings banks; government securities and life insurance. Indeed one aspect of the financial system is much more developed and vastly more important in the Soviet Union than in capitalistic countries, the use of the banking system as a control device for the allocation and use of funds.

An attempt to obtain an over-all picture of the financial system of the Soviet Union for the forty years of its operation, and a comparison of it with the financial system of more or less developed free enterprise and mixed economies, therefore, need no apology. An apology, however, is in order for the very summary fashion in which these comparisons are treated throughout this paper. In particular limitation of the characterization of the financial system of the Soviet Union at different stages of its development, and of the confrontations of the Soviet Union's financial system to a comparison of rough balance sheets for the entire economy with that of Tsarist Russia and the US[2] is due partly to the need of using an approach that provides a maximum amount of relevant quantitative information within a limited space, and partly to the scarcity of serious and statistically founded studies, by Soviet or foreign economists, of the financial system of the Soviet Union.

---

[1] This paper (written late in 1962) is based on material collected in connection with a chapter on an international comparative study of financial structure and growth. It, therefore, should not be regarded as the result of a detailed investigation of Soviet national wealth and finance using all the available Soviet and Western material. The purpose of this paper is limited to bringing out the essential features of the Soviet financial system as they are reflected in the national balance sheet. For that reason, the estimates are only intended to indicate orders of magnitude where no first-hand and trustworthy figures are available. No claim to precision can be made for any of the estimates of financial items until the Soviet authorities themselves publish sufficiently detailed data in this field.

Following an example set in many of Professor Mahalanobis' publications the presentation centers on the hard core — the quantitative data —, and textual comment is kept to a minimum.

II

The financial structure of the USSR is in many respects similar, at least in outward appearance, to that of more or less advanced non-Communist countries, although the similarity seems to have diminished rather than increased over the last forty years.

The USSR has a central bank—the Gosbank—which issues paper money and keeps the government's deposits as central banks in many other countries do, and operates through a nation-wide system of branches and agencies. The Gosbank also supplies virtually all short term credit to productive enterprises and government organizations[3] and keeps the checking accounts of enterprises and the few individuals who have them. Since there are no other banks in the country which provide scriptural money, the Gosbank obviously cannot fulfil one of the chief functions of central banks: the control of the volume of money through the regulation of commercial bank reserves, rediscount policies or operations in the market for government securities. A combination of central and commercial banking functions in the same institution was and is not unknown in other countries. What is specific to the Soviet financial system is, first, the monopoly of a single government institution in the provision of short term credit and the issuance of scriptural money; and, secondly, the far reaching control function of this institution over the financial transactions and even the internal affairs of the enterprises which are bank borrowers and depositors.[4]

The second characteristic of the Soviet financial structure is the combination of short and medium term credit, again on a country-wide almost monopolistic basis in the same institution. This factor is relatively new. In earlier periods of Soviet economic development special institutions, government owned and controlled like the Gosbank, acted as channels for the medium and long term credit needs of some sectors of the economy. These institutions were merged with the Gosbank during the second half of the 1950's with the exception of the Stroibank (former Prombank) which, in addition to serving as a channel of budget funds for investment, supplies medium term credit for construction and the Vneshtorgbank, which attends to the financing of foreign trade.

---

[2] I should have liked to extend the comparison to India but lack of a national balance sheet for India has limited the comparison to occasional references.

[3] The reservation "virtually" is necessary because of the existence of two smaller banks, one for the construction industry (Stroibank) and the other for foreign trade (Vneshtorgbank). The degree of independence of these two banks from Gosbank is in doubt. The Stroibank is also used as the channel through which the non-repayable grants from the budget are made available to government enterprises. These grants which are similar to corporate contributed capital in free enterprise countries, have not been regarded as a financial instrument,

[4] For a description of these relations see Garvy, "The Role of the State Bank in Soviet Planning" in *Soviet Planning: Essays in Honor of Professor Naum Jasny* (to be published in 1963).

As a result of these developments, the Soviet financial structure is now almost monolithic; the Gosbank accounting for over four-fifths of the assets of all financial institutions. The only other financial institutions of importance are the savings banks which have a nation-wide net of branches considerably more numerous than the Gosbank (Table 1). The savings banks (since January 1963 under the jurisdiction of Gosbank), however, like the much smaller single insurance organization (Gosstrakh) are only conduits, all of their assets apart from small cash reserves formerly consisting of 3 per cent Treasury obligations.[4a] This feature of the operation of the savings banks and insurance organizations again is not unique with the USSR, at least not with respect to the savings banks for which the situation is similar in France, the United Kingdom, and some other non-Communist countries. The social security organizations play no role as financial institutions since they operate on a pay-as-you-go basis, and apparently do not accumulate a substantial amount of assets. None of the other types of financial institutions found in most other countries, —such as the building societies (building and loan associations in the United States), credit unions, mortgage banks and investment companies and trusts—is represented in the USSR.

Simplicity in institutional structure is matched by the limitation of types of financial assets in existence. There are, of course, no corporate securities—bonds or stocks. Paper currency, plus relatively small amounts of coin, are the exclusive means of payment in retail trade and for personal services and taxes. Although the Gosbank has 4.2 million demand deposit accounts, apparently only a few belong to individuals. (Individuals, in addition to currency, can use the post office and savings banks for making out-of-town payments.) On the other hand, virtually all transactions among enterprises and between enterprises and the government are handled by cheques or book transfer in the accounts kept with the Gosbank. In this field the use of currency is limited by law to very small transactions to ensure control by the Gosbank over the bulk of financial transactions. Since the Gosbank has not published a balance sheet for more than twenty years, the volume of currency and scriptural money is not exactly known and it is difficult to calculate the velocity of money. Rough estimates suggest that the volume of currency and enterprise demand deposits is considerably smaller in relation to gross national product (together in the order of one-eighth to one-sixth) than is the case not only in financially developed but also in less developed countries. The velocity of money therefore is apparently high, particularly if government and savings bank deposits with the Gosbank are excluded, which account for a substantial proportion of that institution's total deposits.

There is some uncertainty about the extent of accounts receivable and payable. In principle government owned non-financial enterprises are not permitted to extend credit to each other, thus removing the main source of accounts receivable and payable in other countries. Charges for goods shipped but not yet paid for by the buyer are treated as part of inventory rather than as trade credit. The statistics of the working capital of enterprises, however, show substantial amounts under the headings of "debtors" and "creditors."[5] These probably include accrued wages, social security contributions and taxes and some other items of unknown character.

[4a] These assets are now kept in the form of time deposits with Gosbank. This does not affect the total volume of financial instruments outstanding, but changes their breakdown by form and by debtor compared to that shown in Table 2.

The only financial assets available to individuals, apart from currency, are savings bank deposits, government bonds, and life insurance contracts. All three have been subject to repeated unilateral changes in amounts and terms. Of the three, government bonds are quantitatively most important. Most of them are the result of compulsory subscriptions during the war and immediate post war period. The total amount of government bonds held by the public is equal to approximately one-sixth of a year's gross national product and one-third of a year's personal income. However, since most of the bonds now bear no interest and redemption has been postponed to a distant future date, it is doubtful to what extent these bonds should be regarded as individuals' financial assets or Treasury liabilities.

Savings bank deposits are now the main remaining financial assets of the population other than currency. With almost 40 million accounts most urban families apparently have some money in a savings bank, while the 13 million of rural savings bank deposits represent only a minority of the farm population. The average balance of 200 roubles is modest.[5a] It is interesting to compare savings deposits in the USSR with the deposits with the savings departments of savings banks, mutual savings banks, savings and loan shares and deposits in credit unions and the postal savings system in the United States (Table 1). The number of accounts is much smaller in relation to the population in the USSR with about one-fourth compared to a ratio of two-thirds in the United States.[6] The balances in savings accounts of all types are equal to less than 15 per cent of the year's income in the USSR compared to more than 60 per cent in the United States. Compared to the situation in less developed non-Communist countries, the Soviet ratio, however, probably is not low. In India, for example, savings deposits (time deposits in commercial and cooperative banks and with the Post Office) are equal to only about 10 per cent of national product.

III

Economists who have worked with Soviet statistics will realize the difficulties encountered in an attempt to construct a national balance sheet, including both tangible assets (national wealth) and financial assets, for half a dozen of benchmark years over the past half century. It is true that as far as national wealth is concerned the investigator is probably in a better position for the beginning and the end of the period than he would be for most other countries. For the first date we can use the detailed estimates for 1913 of Professor Vainstein in his study of the national wealth of pre-revolutionary Russia.[7] The Soviet wealth census of 1959 provides data for the closing date of the period that are based on an investigation of a scope and thoroughness hardly duplicated in any other country, although the publication of the resulting figures, unfortunately, has been fragmentary, and the description of the methods and sources used is well below the standard usual in such projects in other countries. For the intervening bench mark dates, i.e., the years 1928, 1937 and 1950, the situation is much less satisfactory. For 1928 it is still possible to use nearly contemporary estimates. The figures for 1937 and 1950, however, have to be constructed by extrapolation of the bench mark estimates for 1928 and 1959. The rough procedures that must be used and are explained in Table 3 provide estimates that remain subject to a considerable margin of error.

---

[5] E.g., *Narodnoie Khoziaistvo SSSR v 1960 godu*, pp. 93—95.

[5a] The figures throughout this article, except in Table I, refer to "old" rubles, each of which is equal to one tenth of a "new" (current) ruble, the latter having a foreign exchange value of $ 1. 11.

[6] Even allowing for the larger extent of duplication among savings accounts in the United States a considerable difference in favour of the United States would remain.

A special problem is created of the fact that the Soviet statistics include investment in unfinished construction projects (which is not limited to structures, but also embraces machinery not yet installed) neither under fixed assets nor under inventories. The amounts involved are quite substantial. At the end of 1959 for instance they came to 160 bill roubles (*Kapitalnoie Stroitelstvo*, p. 126) which is equivalent to not less than 5 percent of total reproducible tangible wealth at that date, and more than 20 percent of inventories. These amounts, therefore, are entirely omitted from the official estimates of reproducible wealth and this practice has been followed in Tables 2 and 3. The omission, however, is large enough that it should be taken into account in the calculation in the capital output ratios and in their comparison with those of other countries.

Additional serious problems arise in connection with two important components of national wealth. The first is consumer durables for which no estimate, official or otherwise, is available for the Soviet period. Worse than that, there are no reliable and comprehensive long time series of the value of output or sales of consumer durables from which an estimate could be derived by the perpetual inventory method. The figures entered for consumer durables are, therefore, nothing more than rough guesses.

In the case of land, including subsoil assets, the difficulties are both statistical and conceptual. No estimate has been published for any date later than 1928, and the estimate for that date turns out to be nothing but the prewar figure. Since under the Soviet regime virtually all land, as well as all subsoil assets, is owned by the Government or by farm cooperatives, there are no transactions on which an estimate of market prices could be based. One is, therefore, forced to resort to the capitalization of land rent, a procedure very unsatisfactory in a situation where the determination of land rent is difficult both conceptually and satistically and where there is grave doubt about the capitalization factor to be used. Rough analogies with countries with agricultural conditions similar to that of the Soviet Union, apart from the absence of private ownership of land, must, therefore, form one of the bases of the rough figures entered primarily for purposes of international comparisons.

---

[7] *Natsionalnoie bogatstvo i narodnokhoziaistvennoie nakoplenie predrevolutsionno: Rossi,* 1960.

The estimation of the financial components of the national balance sheet is conceptually easier, but statistically more difficult than in the case of national wealth. The financial assets that still exist in the Soviet Union all have the form of claims of determinable face value. Hence valuation problems are practically absent. The difficulty rather is that the Soviet authorities have published only very few figures in this field, apparently regarding even information on the amount of money in circulation as a state secret. Nevertheless, most of the main components of financial assets can be estimated within a probably reasonable margin of error. Equities, primarily in the form of corporate stocks, that often raise difficult problems in the compilation of the national balance sheets of other countries, of course, do not exist in the Soviet Union. No reliable information is available on monetary reserves, including the holdings of monetary metals. It is certain, however, that any error in the very rough estimates of this item that must be used cannot seriously affect either wealth and figure for national wealth or national assets, or the structure of national total assets.

IV

The national balance sheet is a tool for analyzing the financial structure of an economy. In developed form it consists, first, of a balance sheet for each of the major sectors of the economy, using uniform (market or replacement cost) valuations for all assets and liabilities; and secondly, a combined national balance sheet which permits the study of financial interrelations among sectors at the balance sheet date. A series of sectorized balance sheets provides a basis for an analysis of changes in financial structure and interrelations. The material is not yet sufficient for a detailed sectorized balance sheet of the USSR over all, or most, of its history. Practically the only date for which a sectorized balance sheet can be drawn up, is the end of 1959, the year of the fixed asset census. Even for that date we have to be satisfied with four broad sectors—nonfarm households; nongovernment agriculture, including both the collective (kolkhoz) and the private sector; financial institutions; and nonfinancial government enterprises including government agriculture (primarily sovkhozes) and general central and local government activities.

Since some of the features of the national balance sheet for 1959 will be discussed in the following sections in connection with a review of the trend in the national balance sheet since 1913 and of the characteristics of the financial structure of the Soviet Union, we may limit ourselves here to a few remarks on the financial interrelations among the four main sectors of the Soviet economy as they appear in Table 2.

If collective farms are regarded as outside the sphere of government ownership and management, it appears that the government (central and local) owns approximately 70 per cent of all national assets, about 20 per cent being attributable to non-government agriculture, mainly kolkhozes, and about 10 per cent to the urban population. However, since kolkhoz members cannot individually or collectively sell the land and equipment which they operate; and since most of the government securities held by the population are frozen, without interest and without definite repayment date, it may be more realistic to limit non-government owned assets to privately owned residences and consumer durables and to private holdings of currency, savings bank deposits and life insurance. In that case privately held assets amount to a little over 15 per cent of all tangibles and 10 per cent of all financial assets in the Soviet Union. Of this total, the urban and the rural private sector each own about one-half.[8]

The financial relations among the four sectors are summarized in the following table (*billion rouble*).

sector		claims		liabilities		net claims	
		A*	B*	A*	B*	A*	B*
non-government							
	urban	295	95	5	5	+290	+90
	rural**	130	55	40	40	+90	+15
government							
	financial	580	580	580	580	—	—
	non-financial	595	595	975	700	-380	-105
all sectors		1600	1325	1600	1325	0	0

*A includes and B excludes frozen government securities.

**Including kolkhozes.

One of the outstanding features of this picture of intersectoral financial relations is the virtual absence of debt of the household sector.[9] Hence financial relations, apart from the public's savings in the form of currency, bank deposits and Government bonds, are essentially limited to the government's ownership of financial institutions on one hand, and to the deposit and credit relationships between government owned non-financial and financial enterprises on the other.

V

The main purposes of the unsectored national balance sheet of the Soviet Union shown in Table 3 are, first, to permit the analysis of changes in the relations between the real infra-structure represented by reproducible and non-reproducible tangible assets and the financial superstructure evidenced by claims and liabilities; and secondly, to make possible the study of the structure, and changes therein, of physical and financial assets. The national balance sheets of the Soviet Union for the benchmark dates of 1928, 1937, 1950 and 1959 together with that of Tsarist Russia for 1913, shown in Table 3, should provide a starting point for such an analysis, notwithstanding all its avoidable and unavoidable shortcomings.

---

[8] This estimate is obtained by allocating all of residential structures and consumer durables and one-half of the livestock in Column (2) of Table 2 to the rural private sector.

[9] This statement would be only slightly affected by inclusion of the small volume of consumer credit extended by retail trade establishments which is not included in the tables.

At the present time, or more exactly at the end of 1959, the value of the total tangible assets of the Soviet Union amounts to nearly three times all its financial assets. Hence, the financial interrelations ratio, as it is usually calculated, is equal to nearly one-third. The ratio appears to have been on approximately the same level in both 1950 and 1937, but to have amounted to not much over one-tenth in 1928. From these figures one would conclude that the relative size of the financial superstructure of the Soviet Union has remained fairly constant during the 1950's after a sharp rise in the 1930's and after its almost complete disappearance during the revolution and the period of war Communism. The financial interrelations ratio of the Soviet Union is now almost back to the level it had reached in Tsarist Russia just before World War I, although the components of both numerator and denominator of the ratio are quite different from what they were fifty years ago.

Because of conceptual and statistical doubts about the validity of the estimates of the value of land, it may be preferable to look at the ratio of financial assets to reproducible tangible wealth only. There then appears a considerable degree of stability from the 1930's on at a value of slightly below one-half. The difference to the situation prevailing in Tsarist Russia, however, is now considerably larger, the present ratio of somewhat less than 0.50 comparing with the 1913 ratio of approximately unity. This fact, as well as the differences in the movements of the two ratios during the Soviet period, reflect the declining share of land in total national wealth, which is a necessary concomitant of the industrialization and urbanization which the Soviet Union has witnessed since the 1920's.

Where does the financial interrelations ratio of the Soviet Union stand in international comparison? The present ratio of about one-third is obviously much lower than the ratios in the neighbourhood or slightly in excess of unity that have been common throughout the twentieth century in Europe and the United States and which has now also been reached in Japan.[10] The difference between the Soviet Union and Western countries is somewhat, but not radically, reduced if the denominator is limited to reproducible tangible wealth. The Soviet ratios of slightly below one-half then compares with the American ratio of about unity since the 1920's.

Probably more significant is the fact that the present financial interrelations ratio of the Soviet Union of close to one-third is at about the same level as it is in less developed non-Communist countries. The Soviet ratio is somewhat below those for the larger Latin American republics, but it is approximately as high as the Indian ratio and certainly is above the ratio for most African and several South East Asian countries. At the moment it is unfortunately not possible to go beyond these vague statements because the ratio can be estimated only for relatively few underdeveloped countries and even for those is affected by a large margin of error.

---

[10] For U.S. ratios see, for instance, R. W. Goldsmith, "Le Bilan National des E.U. depuis la guerre," in *Bulletin d'Information et de Documentation de la Banque Nationale de Belgique*, September, 1960; for selected other countries see R. W. Goldsmith, in *Capital Formation and Economic Growth* (National Bureau of Economic Research, 1955), pp. 149 ff.

The financial interrelations ratio of the Soviet Union thus is somewhat lower than one would expect from its real national product per head or from the level of its technological development. This is hardly astonishing in view of the radical differences in financial organization. Indeed it is possibly more remarkable that if equities are omitted from the numerator of the ratio since they are non-existent in the Soviet Union, the level of the ratio of claims to reproducible tangible wealth of approximately one-half which is observed in the Soviet Union is probably not out of line with that of non-Communist countries in a similar stage of development. This numerical similarity, of course, may not mean much because of the different character of some important types of financial assets in the Soviet Union.

VI

The similarity between the national balance sheet of the Soviet Union and that of non-Communist countries is greater for tangible assets or for reproducible tangible wealth than for aggregate national assets.

The ratio of net national wealth to gross national product, i.e., the average capital-output ratio, as usually defined, of the Soviet Union of about three is somewhat lower than that of the United States, for which it is now in the neighbourhood of 3-3/4; or for India for which it was in excess of 3-1/2 at the date of the latest wealth estimate (1950). The difference is greater if the comparison is limited to reproducible tangible assets. Here the Soviet ratio of 2 is well below the United States ratio of 3, but slightly above the Indian ratio of 1-3/4. The reason, of course, is that the share of land in the Soviet Union is considerably above the American but well below the Indian level.[11] These figures are shown in Table 4 together with comparable data for a few other countries.

The changes in the structure of tangible assets shown in Table 3 reflect the transformation during the last forty years of the Soviet Union from a predominantly agricultural to an industrial country. Before World War I land (including forests) accounted for about three-fifths of total national wealth while reproducible tangible assets contributed the remaining two-fifths. By 1937 the ratio had been reversed, the value of land probably not exceeding two-fifths of national wealth. At the present time, it would be difficult to value land at much more than one-fourth of the total national wealth of the Soviet Union.

Given the difficulties of putting a value on land under Soviet conditions it may be better to limit the analysis to the structure of reproducible tangible wealth alone. According to the 1959 Soviet census and a few supplementary estimates explained in the footnotes to Tables 2 and 3, about 30 per cent of the reproducible tangible wealth of the Soviet Union had the form of housing and durable consumer goods. Most of the remainder was represented by the reproducible assets of enterprises. Nonresidential structures accounted for more than one-fourth of the total of reproducible tangible assets, equipment for about one-eighth, inventories for more than one-fifth and livestock for more than one-twentieth.

---

[11] For a more detailed comparison of the level and structure of tangible assets in the Soviet Union and the United States see N. M. Kaplan in *Economic Trends in the Soviet Union*, A. Bergson and S. Kuznets, eds, (1963), pp. 121 ff.

This distribution is, strangely enough, not too different from that prevailing in prewar Russia and during the early Soviet period, as can be seen from Table 5. The outstanding change in the structure of reproducible tangible wealth is the sharp decline in the share of livestock from more than one-eighth in the prewar period to not more than 3 per cent at the present time. Substantial changes obviously have occurred in the geographical distribution of reproducible tangible wealth and in its distribution among industries, but the available data are not sufficient to document these changes, detailed estimates usually being limited to the end of the period.

Probably more interesting is the comparison of the present structure of reproducible tangible wealth in the Soviet Union with similar figures for several other types of economies. Such a comparison is given in Table 6 with the United States; with two developed Western European countries, one (West Germany) representing a case of rapid and the other (United Kingdom) a case of slow growth in the stock of capital; Japan as the most advanced non-Western country; three less developed non-Communist countries, India, Columbia, and Honduras, almost the only countries of this type of which wealth estimates are available; and finally Yugoslavia as a representative of a modified system of centralized planning.

Because of the differences in estimation methods employed in the figures for the various countries, conclusion drawn from Table 6 must be seriously qualified. Nevertheless a few characteristics of the structure of reproducible tangible wealth of the Soviet Union in comparison to other countries that appear in Table 6 are probably genuine and significant.

The most striking of these differences is the high ratio of inventories to fixed reproducible assets. With 25 per cent of total reproducible tangible wealth the Soviet ratio is the highest among the nine countries; and is higher than it may be inferred to be from fragmentary material for any other country not covered by the tables. There is, however, no reason to doubt the high share of inventories in reproducible assets derived from the estimates. The basic figures for 1959 come from Soviet statistics that seem to be as trustworthy as figures of this type usually are. Moreover, they are not out of line with relationships which appear in the much rougher estimates that can be made for earlier benchmark dates in the Soviet Union. It is also known from economic discussions that have been going on for many years in the Soviet Union, that inventories are relatively high, one of the main reasons probably being that large, though not necessarily well assorted inventories are needed as a buffer to absorb short term errors in planning, and themselves are the results of such errors. One of the reasons for this high ratio is the fact that a considerable part of inventories is probably valued to include turnover tax, particularly the inventories of consumer goods, while structures and equipment are practically unaffected by this tax. If the inventory value included on the average a tax element of 25 per cent, the adjusted ratio of inventories to total reproducible assets would decline to approximately 20 per cent. Two additional, though probably lesser, reasons for the high ratio of inventories are the inclusion of goods in transit, some of which are included in other countries in accounts receivable; and the apparent practice of Soviet statisticians to include unfinished structures with inventories rather than with fixed assets. These differences, however, may account for at most five percentage points in the share of inventories in national wealth. It is also possible that inventories include some military items further inflating the figures.

Another, and possible equally striking, feature of the structure of reproducible tangible assets is the relatively low share of equipment. In view of the Soviet emphasis on industrialization in general and heavy industry in particular and the rapid rate of increase in equipment investment since the 1920's and again since the end of World War II, one might have expected the Soviet share of equipment in either reproducible tangible wealth or at least in nonconsumer wealth to be relatively high. Table 6 does not show the expected relation. Because of international differences in classification, it may, however, be better to combine non-residential structures and equipment. Even then the Soviet share of slightly less than one-half of total reproducible tangible wealth is somewhat lower than in the United States and considerably lower than in West Germany, Japan or Yugoslavia. Indeed of the countries included in Table 6 the only one showing a lower share for non-residential structures and equipment is India; and the only other country with a ratio close to that of the Soviet Union is the United Kingdom. To some extent, of course, the relatively low share of equipment among Soviet reproducible wealth is a result of the high ratio of inventories.[12] An additional explanation is the trend in price relationships which in general was unfavourable to equipment in comparison to other components of reproducible wealth. In a calculation in constant prices, whatever that may mean in comparisons over long periods, the increase in the share of equipment might by substantial.

The share of residential structure in reproducible tangible assets in the Soviet Union is low, particularly if climatic conditions are taken into account. The low ratio, however, is entirely in accord with what is known about unsatisfactory housing conditions.

VII

The real difference in the national balance sheet of the Soviet Union and of non-Communist countries at an advanced level of economic development appears not in tangible but in financial assets. This is evident from Table 7 which compares the structures of financial assets in the Soviet Union and the USA. in 1959.

---

[12] It may be argued that the more appropriate comparison would be between the present Soviet ratio of inventories to total reproducible assets with the U.S. ratio of 1900 rather than that of 1959. If such a comparison were made the difference between the two countries would indeed be somewhat reduced, but it would by no means disappear since the U.S. inventory ratio in 1900 of 12 per cent is still substantially lower than not only the unadjusted but also the adjusted Soviet ratio for 1959.

The first difference, the much smaller ratio of financial to tangible assets in the Soviet Union, has already been discussed in Section V.

The second difference is the absence in the national balance sheet of the Soviet Union of corporate stock, which accounts for about one-fourth of all financial assets in the United States, and for substantial though smaller fractions in other developed non-Communist countries. This difference, however, would disappear —or, indeed, point in the opposite direction—if for purposes of national accounting the Soviet government were separated from the non-financial enterprises which it owns and the Government's equity in these enterprises were then regarded as similar to corporate stock.[13]

Thirdly, the ratio of long to short term claims is much lower in the Soviet Union than in the US. Short term claims, consisting primarily of short term credits extended by Gosbank and of deposits with Gosbank and savings banks total over 1100 billion roubles. Long term claims include about 40 billion roubles bank credits, medium term by Western standards; the 110 billion roubles of Treasury securities held by the savings banks and Gosstrakh; and, possibly, the frozen 275 billion of Government bonds, originally of 20 years' maturity, held by the public. The ratio of long (or medium) to short term credits then is in the order of two-fifths if the government securities held by the public are included and only one-seventh if they are excluded. This compares with a ratio of about one to one in the US, the average maturity of the claims classified as long term moreover being considerably longer than in the Soviet Union.

Fourthly, the variety of claims is much smaller in the Soviet Union. There is no parallel in the Soviet Union to the obligations, of different maturity and other characteristics, of state and local governments and of corporations; to the different forms of mortgages; and to the short term marketable securities such as Treasury bills, bankers' acceptances, commercial paper, and finance company papers that are found in non-Communist countries.

There is, however, one aspect of financial structure in which the national balance sheet of the Soviet Union shows an interesting similarity with that of the US, and possibly other developed non-Communist countries. This is the financial structure of the non-agricultural non-financial enterprise sector. The figures for the Soviet Union and the USA for the year 1959 will be found in Table 8. In both countries gross reproducible fixed assets account for fully two-thirds of total assets; and the share of net worth is only slightly higher in the Soviet Union than in the USA. Considerable differences, of course, appear in the higher share of inventories in the Soviet Union, already discussed in Section VI; in the smaller share of financial assets; and in the predominance of bank credit and the absence of long term liabilities in the Soviet Union.

---

[13] In this connection it is important to remember—and this point probably should have been stressed earlier—that most of government and government enterprise capital formation is financed by non-repayable investment grants from the state budget, grants which do *not give rise to financial* instruments and that do not appear among financial assets and liabilities in the national balance sheet.

Statistical Tables

TABLE 1. FINANCIAL INSTITUTIONS IN USSR AND USA, 1960

	Banks[1]		Savings institutions[2]		Life insurance[3]		Three main institutions	
	USSR	U.S.A	USSR	U.S.A	USSR	U.S.A	USSR	U.S.A
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. number of units : 000	0.0	13.5	0.0[4]	26.9	0.0	1.4	0.0	41.8
2. number of offices : 000	6.1	24.0	66.5	33.5	.	.	.	.
3. units : per mill.[5] pop.	0	74	0	148	0	8	0	230
4. offices : per mill. pop.	28	132	310	184	.	.	.	.
5. assets, total : billion[6]	48	311	11	119	1	120	60	550
6. assets : per mill. pop.[6]	223	1710	51	654	5	660	279	3024
7. assets, p.c. of GNP[7]	27.3	61.7	6.3	23.6	.6	23.8	34.3	109.1
8. number of accounts: million[8]	4	120[9]	52	60	10	118	66	298
9. accounts: p.c. of pop.	2	66	24	33	5	65	31	164
10. assets per account: 000 units[6]	11.8	2.6[10]	0.2	2.0[10]	0.1	1.0	0.9	1.8

[1] USSR: Gosbank; USA: Federal Reserve and commercial banks.

[2] USSR: savings banks; USA: mutual savings banks; savings and loan associations; credit unions; postal savings.

[3] USSR: Gosstrakh: USA: Life insurance companies.

[4] All savings banks regarded as one enterprise.

[5] USSR: 215 million: USA: 182 million.

[6] New rouble and dollar respectively.

[7] USSR: 175 billion (new) roubles each of which is equal to 10 old roubles ; USA: $504 billion.

[8] Deposit accounts for Cols. (1) to (4); policies for Columns (5) and (6).

[9] Slightly over one half of total are savings accounts, with balances of $67 billion (including time deposits of corporations and foreigners).

[10] If savings deposits in commercial banks are shifted from Col. (2) to Col. (4) the average per account becomes 4.0 in Col. (2) and 1.5 in Col. (4).

Notes on Table 1

Sources : Columns (1), (3), (5) : Lines 2, 5, 8, *Narodnoie Khosiaistvo SSSR v* 1960 *godu* except for Column (1), line 8 (G. Garvy, "The role of the State Bank in Soviet Planning, in *Soviet Planning: Essays in Honor of N. Jasny* (to be published in 1963), and Column (5), line 8 (R. J. Myers, *Transactions of the Society of Actuaries*, 1959, p. 726).

Line 7 based on GNP estimates of 175 billion (new) roubles extrapolated on basis of Table 2.

Columns (2), (4) and (6) : *Statistical Abstract of the US*, 1962 except for Column (2), line 8 which is based on figures in *Federal Reserve Bulletin*, 1962,

TABLE 2. SECTORIZED NATIONAL BALANCE SHEET OF USSR, 1959
(billion (old) roubles)

	non-government		government[1]		total
	urban	rural[2]	financial	non-financial	
	(1)	(2)	(3)	(4)	(5)
1. reproducible tangible assets	300	480	—	2570	3350
1.1. residential structures	180	180	—	360	720
1.2. other structures	—	50	—	970	1020
1.3. equipment	—	50	—	400	450
1.4. livestock	—	100	—	30	130
1.5. inventories	—	20	—	760	780
1.6. consumer durables	120	80	—	—	200
1.7. gold	—	—	—	50	50
2. land (including forests)	30	600	—	570	1200
3. *financial assets*	296	130	580	595	1600
3.1. currency	10	10	—	10	30
3.2. short term claims	79	42	428	585	1133
3.2a. deposits	79	42	—	420[4]	541
3.2b. bank credits	—	—	395	—	395
3.2c. other	—	—	33	165	198
3.3. medium and long term claims	207	78	152	—	437
3.3a. mortgages	—	—	10	—	10
3.3b. securities	200	75	110	—	385
3.3c. other	7	3	32	·	42
3.4. corporate stock	—	—	—	—	—
4. total assets A	625	1210	580	3735	6150
B[3]	(425)	(1135)	(580)	(3735)	(5875)

	non-government		government[1]		total
	urban	rural[2]	financial	non-financial	
	(6)	(7)	(8)	(9)	(10)
5. liabilities	5	40	580	975	1600
5.1. short term	—	11	570[4]	582	1163
5.2. medium and long term	5	29	—	393	437
5.2a. mortgages	5	5	—	—	10
5.2b. treasury obligations	—	—	—	385	385
5.2c. other	·	24	10	8	42
6. net worth A	620	1170	—	2760	4550
B[3]	(420)	(1095)	—	(3035)	(4550)
7. total A	625	1210	580	3735	6150
B[3]	(425)	(1135)	(580)	(3735)	(5875)

---

[1] Includes government enterprises.
[2] Includes cooperative farms (kolkhozes).
[3] Excluding frozen government bonds.
[4] Includes net worth of Gosbank.

Notes on Table 2

Sources : Figures in Columns (5) and (10) of all lines taken or derived from Table 3.

Line 1.1 : Allocation of total based on data in N. M. Kaplan, "Capital Stock" in *Economic Trends in the Soviet Union*, pp. 112-115.

Lines 1.2–1.3 : Estimates for rural sector based on Kaplan (*loc. cit.*). Soviet figures for kolkhoses alone at end of 1961 (*Narodnoie Khosiaistvo SSSR v* 1961 *godu*, pp. 420-25) are about 100 billion for lines 1 and 2, 25 billion for line 3 and 50 billion for line 4.

Line 1.4 : Total split on basis of number of animals.

Line 1.5 : *Narodnoie Khosiaistvo*, 1961, pp. 382-3. Figure in Column (2) is a rough estimate.

Line 1.6 : Allocation of total between urban and rural population based on income relationship. (Cf. for 1955 M. Bornstein, in *Review of Economics and Statistics*, 1962, p. 448).

Line 2 : Value of private urban residential land (Column 1) assumed equal to about 20% of net value of structures. Value of non-government agricultural land (Column 2) estimated in two steps. Value of all farmland obtained by capitalizing rent, estimated following Bergson's technique (*The Real National Income of the Soviet Union*, p. 138) at 40% of labour income of agriculture (of about 300 billion roubles based on Bornstein's estimate of 250 billion roubles for 1955, *loc. cit.*), at 12-1/2%. This yields a figure of 1000 billion roubles. Of this total 60% is allocated to non-government agriculture, based on its share of 56% of sown area (*Narodnoie Khosiaistvo*, 1961, p. 316) slightly, and probably insufficiently, increased on account of on the average poorer quality and location of sovkhos land. To the remainder, i.e., value of government agricultural land, is added (a) government residential land at about 20% of net structure value; (b) other government non-agricultural land at about 10% of net structure value; (c) very rough estimate for forest land to obtain Column (4).

Line 3.1 : Of assumed total currency circulation of 30 billion roubles (based on an unpublished estimate of R. Powell of 22 billion for 1951) one-third each has been allocated arbitrarily to urban, rural and government sectors. This estimate of currency is probably too high.

Line 3.2a : Column (1) represents savings bank deposits of urban population (*Narodnoie Khosiaistvo*, 1960, p. 854). Column (2) includes savings banks deposits of rural population plus kolkhos deposits with Gosbank estimated at 20 billion roubles. Column (3) consists of demand deposits of enterprises and budget and other deposits of Treasury with Gosbank. It is obtained by deducting kolkhos deposits from total assets of Gosbank, and therefore includes Gosbank's net worth which is attributable to Treasury.

Line 3.2b : *Narodnoie Khosiaistvo*, 1960, p. 849.

Line 3.2c : Figure in Column (4) is sum of "kreditori" and one half of "debitori" as shown in *N. K.*, 1960, pp. 92-93, assuming partial overlap between the two series.

Line 3.3a : *Narodnoie Khosiaistvo*, 1960, p. 849.

Line 3.3b : Total of Government securities held by population (Table 3) split among Columns (1) and (2) in proportion of incomes. Column (3) consists of Treasury securities held by savings bank and Gosstrakh, assumed to equal their assets.

Line 3.3c : Life insurance claims (Columns (1) and (2) split) among urban and rural population in proportion of savings deposits. (This is close to ratio of income of two sectors, according to estimates of Bornstein for 1955). Column (2) is liabilities of Gosstrakh, the counter-entry to line 3.3c, Columns (1) and (2). Entry for government (Column 3) represents medium and long term bank credits (*Narodnoie Khosiaistvo*, 1960, p. 849).

Line 5.1 : Column (7) from *Narodnoie Khosiaistvo*, 1960, pp. 94-5; 849 (all credits to agriculture less those to government agriculture). Column (8) includes all liabilities (including net worth) of Gosbank (estimated at 470 billion) and of savings banks. Column (9) is residual and therefore absorbs errors in other liability items. It is, however, compatible with the sum of short term bank credits (line 3.2b less 5.1, Column (7) and of trade credits (line 3.2c, Column (4)).

Line 5.2a : *Narodnoie Khosiaistvo*, 1960, p. 765.

Line 5.2b : Includes 110 billion roubles of Treasury obligations held by savings banks and Gosstrakh.

Line 5.2c : Figures in Column (6) from *Narodnoie Khosiaistvo*, 1960, p. 849.

Line 6 : Estimate A includes and estimate B excludes government securities held by public.

---

* capitalized

TABLE 3. NATIONAL BALANCE SHEET OF THE USSR, 1913–1959

(billion (old) roubles)

end of	1913		1928	1937	1950	1959
	Empire	USSR				
	(1)	(2)	(3)	(4)	(5)	(6)
1. tangible assets	149.0	125.2	213.5	646.0	2370	4550
1.1. reproducible	57.4	45.8	93.5	396.0	1570	3300
1.1a. residential structures	9.0	7.2	22.9	78.0	410	720
1.1b. other structures	18.6	14.5	39.5	155.0	650	1020
1.1c. equipment	7.7	6.3				450
1.1d. livestock	7.5	6.1	7.9	38.0	80	130
1.1e. inventories	9.5	7.6	16.2	85.0	350	780
1.1f. consumer durables	5.0	4.1	7.0	40.0	80	200
1.2. land (including forests)	91.6	79.4	120.0	250.0	800	1200
2. foreign assets (net)	—5.1	—3.9	0.2	...	...	50
3. financial assets	60.0	48.0	20.0	180.0	750	1600
3.1. financial institutions	18.9	15.0				
3.1a. state bank	3.0	...	4.1	57.3	205	480
3.1b. commercial banks	6.2	...	...	...	...	...
3.1c. savings banks	1.7	...	.2	4.5	19	101
3.1d. mortgage banks	4.3	...				
3.1e. insurance	.3	...	...	...	...	10
3.1f. other	3.4	...				
3.2. financial instruments						
3.2a. government bonds	8.8	...	1.4	28.5	175	385
3.2b. corporate bonds	0.5	...	...	...	...	...
3.2c. corporate stock	4.0	...	...	...	...	...
3.2d. mortgages	5.1	...		1.0	6	10
3.2e. bank credit	...	...	8.6	47.7	193	438
3.2f. other credit	...	...	...	...	110	165
4. national assets (1+2+3)	204.3	169.3	233.7	826.0	3120	6150
5. gross national product	22.0	17.5	32.3	280.7	912	1660

Notes on Table 3

Sources : Columns (1) and (2): Lines 1.1 and 2: A. L. Vainshtein, *Narodnoie bogatstvo i narodnokhoziaistvennoie nakoplenie predrevolutsionoi Rossii*, 1960, pp. 368—71. Figures for "clothing and shoes" have been eliminated from 1.1 and 1.1f. Split between 1.1a and 1.1b is a rough estimate (based on Vainshtein's figure for urban residential and commercial structures). Residential structures exclude farm dwellings. Item 1.2 includes orchards. Item 2 in Column (2) has been set at 80 per cent of Column (1) in accordance with average relation for other lines. Foreign balance includes monetary metals (2.2 in Column 1).

Line 3 : In all years an estimate based on figures given below and rough guesses for missing items.

Lines 3.1a–3.1f: *Statistitcheski Yejegodnik*, 1914 and 1915. Line 3.1d includes Nobles' Bank, Peasants' Bank and State Mortgage banks; line 3.1f: mututal credit associations, urban credit associations, urban banks and small credit institutions.

Line 3.2a: Same as line 3.1a–3.1f. Only securities of central government.

Line 3.2b: Rough estimates based on total liabilities of non-financial corporations of 2.7 billion roubles (*op. cit.*)

Line 3.2c: Dividends of non-financial corporations (about .25 billion roubles, *op. cit.*) capitalized at 6 per cent.

Line 3.2d: Estimate of real estate debt as of 1/1/1913 by V.A. Mukoseiev in A. Raffalovich ed., *Russia: Its Trade and Commerce*, p. 387.

Line 4: Katz's figure (*Narodni Dokhod SSSR*, 1932, p. 60) for USSR plus allowance of 25 per cent for omitted items (services; depreciation). Figure for Empire assumed 25 per cent higher.

Column (3): Lines 1.1a–1.1e: *Kontrolnie Tsifri na* 1928–29 *godu*, pp. 426 ff.; S. Rosentul, *Planovoie Khoziaistvo*, 1929, I, pp. 308–09. Figures on prices of 1925–26 refer to October 1.

Line 1.1f: Rough estimate for 1928.

Line 1.2: Estimate of Strumilin (*Ocherki Sovietskoi Ekonomiki*, p. 21) of 59.6 billion roubles for October 1, 1924 at prewar prices (then equal to 150 per cent of reproducible fixed assets according to Strumilin's estimates) increased by about 700 per cent for changes in price level between 1913 and 1928. (*Ekonomitcheski Bulletin*, Koniunkturnyi Institut).

Line 2: *Statistisches Handbuch der Weltwirtschaft* (ed. German Statistical Office), 1936, p. 256.

Line 3.1a: R. P. Powell, unpublished estimates.

Line 3.1c: R. P. Powell, *Soviet Monetary Policy* (unpublished, Ph.D. Thesis, University of California), p. 87. Figures refer to individuals' deposits; in addition, deposits of organizations of 0.1 in 1928 and 0.2 in 1937.

Line 3.2a: R. P. Powell, *Soviet Monetary Policy*, p. 78.

Line 3.2b: Ia Golov, in *Sovietskie finansy sa XXX let*, 1947, pp. 110–11 and 126.

Line 5: Estimates of Bergson, *The Real National Income of Soviet Russia*, p. 46.

Column (4): Line 1.1a–1.1d: Rough estimates based on changes in value compared to 1928 in 1955 prices (*Narodnoie Khosiaistvo*, 1959, p. 66; 1937–40 increase assumed proportional to increase in national income, *op. cit.*, p. 77) and in prices (Bergson's deflators as derived from *The Real National Income of Soviet Russia*, pp. 46, 48).

Line 1.1e–1.1f: Rough estimate based on 1928 relations.

Line 1.2: Bergson's estimates of agricultural land rent of 25 billion roubles. (*The Real National Income of Soviet Russia*, p. 138) capitalized at 12-1/2%, plus rough estimates for other land.

Line 3.1a, 3.1c: Same sources as for 1928.

Line 3.2a, 3.2b: Same sources as for 1928.

Line 3.2d: Estimated on basis of 1940 figures of 1.2 billion roubles (*Narodnoie Khosiaistvo*, 1960, pp. 852–3). Includes only loans on individually owned homes.

Line 5: Same sources as for 1928.

Columns (5) and (6): Lines 1.1a–1.1c: 1950: Estimated at 53% of value of 1959, on basis of quantity relation (49%) (*Narodnoie Khosiaistvo*, 1959, p. 66) and 8% decline in prices between 1950 and 1955, based on Bergson's implicit deflators (*op. cit.*, p. 85). 1959: Soviet Census figures of gross stock of capital (*Narodnoie Khosiaistvo*, 1960, pp. 88 ff.) reduced by percentage of "wear and tear" indicated in census, i.e., 23 per cent for housing, 22 per cent for other structures and 27 per cent for equipment. Figures are expressed in prices of 1955, which probably do not differ considerably from those of 1959.

Line 1.1d: 1950: Rough estimate; 1959; N. M. Kaplan, in *Economic Trends in the Soviet Union*, 1963, pp. 111 ff.

Line 1.1e: Estimated on basis of figures in *Narodnoie Khosiaistvo*, 1960, pp. 92 ff. which do not include collective farms, but do include immature livestock.

Line 1.1f: Estimated on basis of average life of consumer durables of 15 years. Volume of sales of consumer durables for 1955, 1958, 1959 from *Narodnoie Khosiaistvo*, 1960, p. 69; for other years estimated on basis of growth rate as given in R. E. Golden, "Recent Trends in Soviet Personal Income and Consumption," in Joint Economic Committee, *Dimensions of Soviet Economic Power*, p. 360. The 1950 figure is a rough guess making allowance for low level of production of consumer durables during 1940's and war losses.

Line 1.2: 1950; Assumed at 48% of line 1.1, the average of the ratios for 1937 and 1959. 1959: See Table 2, line 2.

Line 2: Only monetary gold stock. Estimates based on data in O. Altman, *IMF Staff Papers*, VII, pp. 421 ff. (1959 value is midpoint of Altman's 1958 range).

Line 3.1a: Based on unpublished estimates of R. P. Powell.

Lines 3.1c, 3.2d, 3.2e: *Narodnoie Khosiaistvo*, 1956, p. 282; 1959, p. 807; 1960, p. 849. Line 3.2d is limited to loans on individually owned homes.

Line 3.1e: Based on estimate of face value of life insurance of 50 billion roubles (R. J. Myers, in *Transactions Society of Actuaries*, XI, p. 726) and ratio of reserves to face value of 20 per cent.

Line 3.2a: 1950: G. Grossman in B. Beckhart, ed., *Banking Systems*, p. 752 for 1951. 1959: Bonds sold to public as of 4/1/57 (267 billion roubles, according to A. G. Zverev, *Gosudarstvenoie zaimi i vkladi*, 1957, pp. 23-26) plus estimated sales to end of 1957 when sales were discontinued. Figures include about 110 billion rouble bonds that have been issued to savings banks and insurance organizations.

Line 3.2b: *Narodnoie Khosiaistvo*, 1960, pp 92-3 ("Creditors" plus one-half of "Debtors").

Line 5: Estimates of A. Bergson for 1950 (*The Real National Income of Soviet Russia*, p. 300). Estimate for 1959 is based on Bergson's figure for 1955 (1184 billion roubles, *loc cit.*), N. Nimitz' estimates for 1956-58 (Rand Memorandum, June 1962); and Soviet estimates of change in national income in 1959 (*Narodnoie Khosiaistvo*, 1960, p. 152) ignoring possible small price change between 1959 and 1960. The absolute figure given for 1959 is 1360 billion roubles (*op. cit.* p. 153), on Soviet definition of national income.

TABLE 4. THE AVERAGE CAPITAL-OUTPUT RATIO IN USSR AND TEN SELECTED COUNTRIES
(multiple of gross national product)

country	year	total national wealth		reproducible tangible assets	
		including	excluding consumer	including durables	excluding
(1)	(2)	(3)	(4)	(5)	(6)
1. USSR	1959	2.8	2.7	2.0	1.9
2. USA	1958	3.8	3.4	3.1	2.7
3. India	1950	...	3.5	...	1.7
4. Yugoslavia	1953	5.7	5.2	4.2	3.7
5. France	1954	4.1	3.7	3.3	2.9
6. West Germany	1955	2.8	2.5	2.3	2.0
7. Netherlands	1952	2.8	2.5	2.4	2.1
8. Sweden	1952	6.1	...	5.5	...
9. Canada	1955	...	...	2.2	1.8
10. Australia	1956	4.0	3.7	3.3	3.0
11. Japan	1955	...	...	2.5	2.0

Sources:

Wealth: Line 1: Table 3; line 2: R. W. Goldsmith, *The National Wealth of the US in the Postwar Period*, pp. 117; lines 3-11 *Income and Wealth*, Series VIII, pp. 8.

Gross National Product: line 1 and Table 3; lines 2, 4-11 United Nations, *Yearbook of National Accounts Statistics*, 1961; line 3: United Nations, *Statistics of National Income and Expenditures*, 1955.

TABLE 5. CHANGES IN STRUCTURES OF TANGIBLE WEALTH OF USSR

	percent of total tangible wealth					percent of reproducible assets				
	1913	1928	1937	1950	1959	1913	1928	1937	1950	1959
1. residential structures	6[1]	11	12	17	16	15	25	20	26	21
2. other structures	13	18	24	27	23	31	42	40	41	30
3. equipment	5				10	13				14
4. livestock	5	4	6	4	3	13	8	10	5	4
5. inventories	6	8	13	15	17	16	17	20	22	23
6. consumer durables	3	3	6	3	4	8	8	10	5	6
7. gold	1	0	0	0	1	4	0	0	0	2
8. land (including forests)	61	56	39	34	26	154	128	62	51	36

[1] Does not include farm residences, which presumably are included in line 2.

Source : Table 3.

TABLE 6. STRUCTURE OF REPRODUCIBLE ASSETS IN USSR AND SELECTED COUNTRIES

(percent of reproducible tangible assets excluding consumer durables)

	USSR (1959)	USA (1958)	India (1950)	United Kingdom (1953)	West Germany (1955)	Yugoslavia (1953)	Japan (1955)	Columbia (1953)	Honduras (1955)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. structures	56	71	55	...	62	67	57	66	58
1.1. residential	23	35	26	34	29	25	17	16	23
1.2 other	33	36	29	48	33	42	40	50	35
2. equipment	15	18	14		26	15	22	20	15
3. livestock	4	2	16	2	12	4	1	14	38
4. inventories	25	10	16	17		14	21	...	
5. consumer durables	6	16	...	...	17	14	25	...	10
6. net foreign assets	2[1]	5	...	...	...	-1	1	...	-12

[1] Gold only.

Source : Column (1) : Table 3. Column (2) : R. W. Goldsmith, *The National Wealth of the US in the Postwar Period*, p. 117. Columns (3-8) : *Income and Wealth Series VIII* (1959), p. 8. Column (9) : Unpublished estimates of E. Tosco.

TABLE 7. THE STRUCTURE OF FINANCIAL ASSETS IN USSR AND USA, 1959

	absolute figures (billion)		distribution (percent)		percent of gross national product		percent of national wealth	
	USSR	USA	USSR	USA	USSR	USA	USSR	USA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1. money[1]	300	173	19	9	18	36	7	10
2. savings deposits	101	162	6	9	6	34	2	9
3. bank credits, short	395	71[2]	25	4	24	15	9	4
4. bank credits, long	42	...	3	...	3	...	1	...
5. consumer credit	...	52	...	3	...	11	...	3
6. trade credit	165	85	10	5	10	18	4	5
7. treasury obligation, short	...	77	...	4	...	16	...	4
8. treasury obligation, long	385	164[3]	24	9	23	34	8	9
9. state and local govt. securities	...	64	...	4	...	13	...	4
10. corporate foreign bonds	...	85	...	5	...	18	...	5
11. obligation of insurance organizations	10	161	1	9	1	34	0	9
12. mortgages	10	191	1	10	1	40	0	11
13. claims[4]	1600	1371	100	75	97	287	35	76
14. corporate stock	...	455	...	25	...	94	...	25
15. all financial assets	1600	1826	100	100	97	381	35	101

[1] Currency and demand deposits

[2] Excludes consumer credits, but includes all security credits.

[3] Does not include Treasury obligations held by social security organization ($21 billion).

[4] Includes miscellaneous claims

TABLE 8. COMPARISON OF FINANCIAL STRUCTURE OF NON-FINANCIAL NON-AGRICULTURAL ENTERPRISES, USSR AND USA, 1959
(percent)

	USSR	USA
	(1)	(2)
1. reproducible fixed assets, gross[1]	68	67
2. inventories	28	11
3. financial assets	4	22
a. cash	2	5
b. assets receivable	2	10
c. other	.	7
4. liabilities	19	25
a. bank credit	14	4
b. accounts payable	5	7
c. other short	...	2
d. long	...	12
5. net worth[2]	81	75
6. total	100	100

[1] If net value of reproducible fixed assets is used, share is reduced to 62 per cent in USSR and 53 per cent in USA, and share of other assets increased proportionately.

[2] If net value of reproducible fixed assets is used share is reduced to 72 per cent in USSR and 64 per cent in USA.

Notes on Table 7

Sources : USSR: Table 2 and *Narodnoie Khoziaisvto SSSR v* 1960 *godu*, pp. 849 ff.

USA: *Federal Reserve Bulletin*, 1960, p. 939–40. (National wealth estimated on basis of R. W. Goldsmith's figure for 1958 in *The National Wealth of the US in the Postwar Period*, p. 117 ff.)

Notes on Table 8

Sources : Column (1) : *N. K.*, 1960, pp. 86, 93/93.

Column (2) : *Survey of Current Business*, November 1962, pp. 11, 15 (for line 1); Federal Reserve Board, *Flow of Funds/Saving Accounts*, 1946–1960, *Supplement* 5, (for lines 2–4).

ELEMENTS OF A THEORY OF PLANNING

By NATHAN KEYFITZ

Université de Montreal, Canada

SUMMARY

Our argument begins with an examination of the respects in which the planners have a social location and economic preferences different from those of the mass of citizens in the under-developed country, two aspects of which are important; that the planners are part of the urban elite rather than the rural aristocracy, and that even as compared with the mass of city-dwellers in their countries they take a long-range view of problems. We repudiate any view that would locate the planners outside of society and make them all-powerful, either gods or demons, below whom the society is an inert body to be subjected to manipulation. We address ourselves to the problem of how and in what degree the planners can impose their preferences. The activation of people by appeal to self-interest seems the simplest method, but even in the presently developed countries most committed to free enterprise it is used alongside an appeal to duty or altruism, and reinforced by some measure of coercion. In the underdeveloped country, as we shall show below, these three forces are more or less out of control and pull the citizen in opposing directions, so that in consequence his net movement is small. An alignment of the three seems indispensable to development. But as soon as one attempts to guide the process consciously by working with egoism, duty, and coercion, it turns out that these apparently primitive psychological elements are in their turn reflections of social groupings; if altrusim is the internalization of the interest of the nation as a whole, egoism is the internalization of the interest of some sub-group of the nation. The developing country is already rich in such sub-group. Some of these are traditional—ethnic, caste, village, clan; while others are modern forms of association—professional, industrial, trade-union, military. The condition of development is not—as is sometimes thought—that these sub-groups be weakened or destroyed in the interest of strengthening the nation, but rather that none of them be allowed to stand in the way of the development of others. During the course of development there is a steady increase of goods and an elaboration of certain social structures. Included among the latter are those groups devoted to physical production of the entities which the planners want; these must be strengthened in the competition both against less efficient producers and against non-productive organizations. Organizations capable of efficient production must be enabled to rise above the rest of the social landscape, and to devise the rules by which this may happen is the principal intellectual task of the planners; their main administrative task is to provide the capital means to such organizations. The problem of planning is that of the allocation of social power—i.e., capital in the broadest sense—to such corporate entities in order that they may grow to the extent required by the physical needs of development.

The organizations so favoured include steel mills, banks, universities, research institutions. It is essential that these be allowed to develop in the maximum freedom from rules imposed by outsiders; that the condition for their access to capital for expansion be exclusively their ability to attain the goal set for them. This may variously be the output of steel, the efficient lending of funds to producers, the output of genuinely educated graduates, the discovery of new knowledge. The different ways in which the growth of such institutions has been favoured by the presently developed countries provide instruction on how this can be achieved elsewhere.

PLANNING IS POSSIBLE

The possibility of economic planning is now generally conceded. Those who prefer free enterprise, including the present writer, agree with those who prefer socialism that a modern economy can be conducted on either arrangement. "So far as its pure logic is concerned the socialist plan makes sense and cannot be disposed of on the ground that it would necessarily spell chaos, waste, or irrationality."[1] What the computing mechanism of the market is able to attain, a set of prices within which the action of all the individuals in the market is determinate and in some sense rational, can be attained also by planners through the solution of sets of simultaneous equations. "We thus derive demand functions for consumers' goods and supply functions of labour and saving, and ... guided by these functions and by its own technological knowledge the ministry [of production] will cause appropriate quantities of consumers' and investment goods to be produced."[2] It is ironic that this proposition was not discovered by socialists but by their opponents, and above all by Pareto and Barone.[3]

The importance of this theoretical result is less now than when it was first put forth, since there are instances of industrialization before our eyes of all shades and colours: free enterprise in England and America; paternalistic-military in Japan (with the military aspect dropped after World War II); étatist in Turkey; cooperative socialist in Sweden; entirely conducted by the state in the USSR. There are, of course, differences in the rates of progress shown by countries, but there is no evident correlation between the auspices under which the economy is conducted and the rate of its growth. Taking two opposite extremes, W. W. Rostow compares the United States and the USSR and shows their parallelism, with a lag of three and a half decades for the USSR.[4]

It is useful to examine a cross-section of data on the several countries of the world in order to reinforce the point that the degree of conscious planning, or one kind of planning against another, does not account decisively for differences in rates of growth. The unequal results of planning in different countries is a fact whose establishment need not detain us for long. The United Nations has dedicated the most recent of its reviews of world economic conditions to industrialization and economic development.[5] Its compilation of the results of the efforts made in the 1950's is unavoidably based on national statistics, and comparison of these suffers from the handicap that the industrially underdeveloped countries are also those which are statistically underdeveloped; nonetheless it can be assumed that the United Nations has secured the most careful compilations available. During the 1950's the average annual rate of growth of manufacturing output was 5% for India, compared with 1% for Argentina.[6] It is true that India planned more seriously than did Argentina. But Brazil's increase was 7% and it planned less seriously than India. The Federal Republic of Germany took a good deal of pride in not planning at all and rose by an average of 7.6% per year, exceeded by Japan which is also a non-planner. The USSR, which virtually invented planning, claims 11.6% per annum increase in industrial output.[7] And the United States, which believes in planning for underdeveloped countries but not for itself, increased at 3.6% per annum.[8]

---

[1] Joseph A. Schumpeter, *History of Economic Analysis*, Oxford University Press, 1954, New York: p. 989.

[2] *Ibid.*, p. 988.

[3] *Ibid.*, p. 987.

[4] *The Stages of Economic Growth, A Non-Communist Manifesto*, Cambridge University Press, 1960, Chapter 7, "Russian and American Growth, A Remarkable Parallel," pp. 93-105.

PREFERENCES OF THE PLANNER

All planners have in common, in contrast to their constituents, a long view of economic matters. This is one uniformity in a host of differences; it applies to such earliest planners as the conservationists in the United States, who directed their attention to the physical resources that would exist in a future beyond what other people cared about; it applies to Soviet and Indonesian planners alike — to take extremes in the application of effective coercion — in comparison with the perspectives of the mass of the people who are to be the subjects of their plans.

In part the long view arises simply because the planner sees himself as a member of a profession: in all fields the professional prides himself not so much on doing a particular job as on applying to it a technique that has a future and whose custody is his career. He uses the individual case as a means to improve his technique. The surgeon's concern is not for his patient as such but for bettering his precision. Paradoxically it is the surgeon's devotion to the long-term improvement of his own skill that will save the patient's life, not any affection for the patient. The economy likewise benefits more from the planner's detached pride in his craft than from his love of his fellow citizens.

But the long view is imposed also by the technical conditions of modern industry which needs capital, and the simultaneous fact that urban and highly capitalized industrial production is the only means by which the urban elite as a whole, a collectivity with which the planner is invariably identified, can maintain its position within the country. In this the motives of the planner will overlap with those of the entrepreneur in the early phases of a system of free enterprise; both are in effect in competition with rural folk who are spinning, weaving, beating out iron knives. Ultimately somehow the capital accumulated in the city will win over rural methods; the displaced peasants, of low productivity in the countryside, will come to the city and take higher-productivity jobs. In the early stages it is perfectly permissible to contract rural artisans in a system of home industry—found even today in countries as advanced as Switzerland and Japan—but this is not typical of the planner's effort. To secure a firm position based on decisively higher productivity he needs to command overwhelming amounts of capital in the form of buildings and equipment.

---

[5] *World Economic Survey*, United Nations, 1962.

[6] *Ibid.*, p. 36.

[7] *Ibid.*, p. 97.

[8] *Ibid.*, p. 63.

It follows from this that the planner's need is for low rates of interest. If people are ascetic and live well within their earnings, this condition will be met. Limits were set on consumption in England, Japan, and the USSR in their crucial periods of capital accumulation, either by a puritan ethic, or by restriction of purchasing, or both. Asceticism, like some other essentials of development, does not always offer an attractive exterior; the close-fisted French peasant who put his savings into government and industrial bonds was nevertheless an important agent of development. People who are open-handed in their personal affairs do not lend to the planner but compete for resources with him. The competition can easily force interest rates up as high as 50% or more per annum. Most governments, even those not well inclined towards planning, have laws against usury on the books; they want the industry of their country to have the use of the money rather than people who would apply it for consumption. A running competition is going on everywhere in Asia between the planner and the peasantry, characterized by shorter time-preferences on the part of the peasants and constant effort on the part of the planners to build institutional fences—for instance savings banks—to contain any resources which might be saved, so that these may be channelled into production rather than allowed to slip back into consumption.

THE PLANNER'S CAPITAL

It is fortunate that money capital and the real resources which it can command is but a small part of the assets of which the planner can dispose. Homans' definition of capital suggests a vast extension; for him capital is anything that allows one to "postpone actions leading to some immediate reward in order to undertake others whose rewards, though potentially greater, are both uncertain and deferred."[9] Money satisfies this definition in the degree in which the currency can be trusted. A stock of food permits it also, since by means of it one could attend to work of capital construction undisturbed by the need to sow and harvest the grain for one's bread. Adam Smith was very clear on this point; he says of the producer under a regime of division of labour that "a stock of goods of different kinds must be stored up somewhere sufficient to maintain him, and to supply him with the materials and tools of his work," until such time as he completes his product and sells it.[10] But in the underdeveloped country it is not the modernizing elite which controls stocks of food but rather the peasantry. The fact that the food—which constitutes at once the chief consumers' good and the bulk of working capital—is not initially in the possession of the planners, or of the urban elite, but in the storage bins of the peasantry or of their landlords enables us to identify a main political and social problem of planning in any country which is at the start of development.

---

[9] George C. Homans, *Social Behaviour : Its Elementary Forms*, Harcourt, Brace and World, 1961, 86.

[10] *The Wealth of Nations*, Modern Library, 1937, p. 259.

Whatever means will permit the urban elite to gain control of this food, and then to issue it to people on suitable conditions, constitutes capital in our wide sense. The encadrement of India by the Indian Civil Service in the 19th century, like that of the USSR by the Communist Party in the 20th, along with whatever apparatus of law legitimized authority and so made it more effective, constitute such capital. The organization to secure deliveries of food from peasants was supplemented in the USSR by non-peasant production of food on collective farms; the use of tractors and other machines permitted a diminution of the application of force to the human factor. Tractors are a partial substitute for encadrement.

The essence of the several forms of capital is that they are transformable one into another, and in the successive transformations lie the possibilities of net gain. The organizational network, encadrement, can be transformed into stocks of food delivered into urban warehouses; these stocks in turn can be offered to workmen who will build the brick walls of a factory, or they can be sold abroad in order that the factory may be equipped with foreign machinery. There is much room for research as to the means by which the urban elites of the developing countries have been able to extend their power in the interest of development, and the successive conversions of power from one form into another by which it is steadily increased. For the change from organization to food and from food to construction is only a beginning. The factories once built can produce a stream of goods; these can include such consumers' goods as shoes, clothing, and bicycles; consumers' goods strengthen the organization in the measure in which they are allotted to the cadres. The cadres would thereby by spurred to put more pressure on their peasants to secure more of the next crop. But an alternative cycle of transformation is to use the goods directly to reward the peasants for their productive efforts—in other words, encourage a free market whose functioning would gradually dispense with the need for the encadrement. This cycle is, of course, illustrative only; it is presumably one of many which the elite would employ as their means of planning development.

Our liberal upbringing must not cause us to forget that power is an unavoidable constituent of the process. For power is "the generalized capacity to mobilize resources in the interest of attainment of a system goal [and] the generation and utilization of power constitutes one of the fundamental functional imperatives of any social system."[11,12]

---

[11] Talcott Parsons, "Sociological Approach to the Theory of Organizations, II," *Administrative Science Quarterly*, Vol. I, No. 2, September 1956, p. 226.

[12] Frank H. Knight, The Economic Order and Religion, Kegan Paul, 1948, p. 48

THE PLANNER'S INSTRUMENTS FOR ACTING ON THE INDIVIDUAL

This description in terms of the diverse forms of capital is in one respect too abstract. How do the process and the organizations gear into the individual?

I suggest that for present purposes we need to recognize only three ways of getting people to do things: *self-interest*, *duty*, and *force*. We can offer a person rewards, we can appeal to his conscience, we can threaten him with punishment. These are slightly different from Talcott Parsons' coercion, inducement, and "therapy"[13] and less detailed than a list due to Robert K. Merton,[14] but they will serve our purpose.

Self-interest. A large library which includes some of the most solid work in social science has been devoted to the ways in which the pursuit of ego's self-interest, as interpreted by him, can converge with the self-interest of alter and tertius who are apparently in opposition or competition. The resulting interplay ensures not only that the things which need doing by society get done, but that they get done with the allocation of the smallest amount of resources in a kind of automatically computed optimum. The economizing process is so sure and effective that no one needs to plan it in detail or even be aware of its existence. It ensures stable progress whose high degree of continuity and certainty is the result of tying the economy, as Adam Smith said, not to the noblest but to the strongest motives of men. We depend for our bread on the self-interest of the baker, not on his benevolence. The workability of such an arrangement, however, depends on certain indispensable items of pre-existing culture, and no one has established exactly what they are, let alone how widely distributed. They certainly include a framework of formal law (rather than substantive justice) among their preconditions. The job of government is to ensure that as little profit as possible is secured by means other than effective production. Hence laws against gambling, laws which require the railways to carry the goods of all individuals at the same charges, laws against monopoly, laws which grant property in inventions.

The Soviet planners have made constant use of the conception of man as an animal who will maximize his interests, though they have by no means assumed that the efforts of individuals will aggregate in such a fashion as to suit the ends either of the regime or of the nation as a whole. Incentives apply in the domain of consumption as in that of work.[15] At the lower levels it appears that the average worker actually receives about double his standard wage in all, the excess being progressive piece rates for over-fulfilment of his work norm.[16] Alec Nove studied the differentials between skilled and unskilled workers in the same industry and found these to be much greater in the USSR than in western countries : about 2.8 : 1 on the average, compared with 1.3 : 1 in England.[17] Though Stakhanovism is largely gone, the mining industry where it started and where incentives are badly needed shows a differential of 3.2 : 1. The ideal of Socialist competition remains in all industries.[18]

---

[13] "Sociological Approach to the Theory of Organizations, I," *Administrative Science Quarterly*, Vol. I, No. 2, June 1956, p. 79.

[14] "Patterns of Influence," in Paul F. Lazarsfeld and Frank N. Staanton (Eds.), *Communications Research*, 1948-9, Harper and Bros., 1949, p. 218.

[15] R. W. Davies, *Development of the Soviet Budgetary System*, Cambridge University Press, 1958, p. 287.

[16] Charles R. Walker, *Modern Technology and Civilization*, McGraw-Hill, 1962, p. 394.

Socialist incentives frame the activities of the industrial manager ; here again pay increases and bonuses are less automatic and more dependent on personal effort in the Soviet Union than in western countries. Joseph Berliner cites figures for 1954 showing that in the iron and steel industry managerial personnel earned bonuses averaging 51% of their basic income, while at the low end, in the food industry, the average was 21%.

> If a young person decides to find employment in the electrical machinery industry rather than in the furniture industry, the difference in basic salaries in the two industries may well have been the decisive consideration. ...But once in the job, the salary does not vary according to whether one operating decision is made rather than another. When the manager decides to put one order into production ahead of another, or to substitute one material for another, it is not his salary he is thinking about. It is usually the size of the month's bonus that will depend on the decisions taken.[19]

The Soviet planners are if anything over-eager to frame prices of materials and of final products in such fashion that the "profits" which will accrue to managers will be an incentive to them to do what the plan requires. They load so many incentives into the system that people find it hard to calculate the advantages for themselves which the planners believed they were embodying. If the manager of the diesel plant Spirit of the Revolution gets so much for each increment of total production, and so much for each increment of profit, and so much for having no accidents in the plant, and so much for surpassing some measure of quality, he loses track of where his interest lies ; the signals which the planners are trying to send him become attenuated. Sometimes they signal the wrong thing. Berliner tells us that "the Commissariat of the Iron and Steel Industry...once introduced a system in which managers could earn a bonus of 50% of the value of fuel saved and 15% of the value of materials saved. The system may well have turned a good deal of effort away from materials to fuel."[20] At best the incentives require to be reached by "storming" at the end of each month. But all these defects become known and periodic efforts are made to simplify the signals.

If money is the principal signalling system in any modern economy, it is not the only one. Planners can show their gratitude by differentiated uniforms, by medals, by public receptions for especially ardent workers and managers and public ceremonies in their honour. Moore speaks of 19 grades of authority and prestige marked out by uniforms in the coal industry.[21] The advantage to the economy of uniforms and medals over payments in money with which real goods may be bought is very clear : consumers' goods require allocations of equipment and labour which inevitably reduce output of producers' goods. Automobiles made to pay labour reduce to that extent the output of machine tools. Walter Galenson reports that "the Pakistani worker...is as much motivated by traditional values of prestige and honour as by monetary incentives,"[22] and draws the same conclusion : if the non-pecuniary incentives can be properly utilized they may mitigate the severe impact caused by introducing a pure wage-incentive system into an economy not yet geared to the generation of a sufficient quantity of consumers' goods. A means of heightening every form of stimulus—the monetary as the honorific—is the cultural separation of the holders of high status from the mass of the population, with some tendency for the different styles of life to constitute separate marriage groups.[23]

---

[17] "The Worker and His Livelihood," in Alex Inkeles and Kent Geiger (Eds.), *Soviet Society*, Houghton Mifflin, 1961, p. 386.

[18] *Ibid.*

[19] "The Situation of Plant Managers," in Alex Inkeles and Kent Geiger (Eds.), *Soviet Society*, Houghton Mifflin, 1961, p. 366.

[20] Joseph S. Berliner, "A Problem in Soviet Business Administration," *Administrative Science Quarterly*, Vol. I, No. 1, June 1956, p. 100.

In practice self-interest is never the interest of the single individual, but rather the interest of a sub-group of the group in question.[24] At the very least the man who is said to be working for himself is aiming to improve the financial and social position of his family, but even this is in general too narrow a statement. The profit position of a firm—in the United States or the USSR—is in the interest of the managers in a degree which would not be explained wholly by reference to the part of the profits which is paid to them personally. The fact that they can arrange for an output of higher value than the input they use is what gives the collectivity of the firm the claim on society for the resources by which it may continue in operation, whether under capitalism or communism. Every member of the concern wants it to have a strong claim to the resources of society. Managers are so selected as to be especially sensitive to the interest of their firm, and this sensitivity is increased by the influence of the people they meet in their daily round. The manager of the Soviet enterprise that over-fulfils its quota has the right to place a certain portion of his profits in a fund usable at his discretion to provide certain kinds of facilities to his employees—housing, paid vacations, etc. The disposition of such funds directly strengthens his position with his employees and makes it easier for him to call upon them for extra effort the next time storming is required to attain the monthly target. The atainment of sensitivity to the profit motive is part of the process of development, perhaps the main part.[25]

Altruism or goodwill. However, even those who are not too squeamish to hang their social theory on the strong—which are broader than egotistical—motives of men often consider that the use of these motives is ineffectual : it brings too slow a response, or the wrong response, among men not educated to their own interests.

---

[21] Barrington Moore, *Soviet Politics—The Dilemma of Power*, Harvard University Press, 1959, p. 239.

[22] *Labor in Developing Economies*, University of California Press, 1962, p. 9.

[23] Barrington Moore, *op. cit.*, p. 245 ; Robert Bordaz, *La Nouvelle économie soviétique*, 1953-1960, Paris : Grasset, 1960, p. 145.

[24] Talcott Parsons, *Structure and Process in Modern Societies*, Free Press, 1959, p. 114.

[25] Talcott Parsons, *The Social System*, Free Press, 1951, p. 243.

And education takes time. Why not go about the matter more directly ? Everybody wants development, it is assumed; why not take advantage of this common interest to secure the necessary capital accumulation and toil ? We will all work together to secure what all of us want. The function of the plan is to say what it is that each of us will inevitably have to do sooner or later if development is to be secured, and to urge us to do it quickly. The plan speaks also of aggregate consumers' goods production and so is a promise of reward in the form of a higher average income at the end of the plan period, but the reward is a collective one, no specific shares being promised to individuals. Full advantage is taken of modern means of communication which bring the collective tasks to the attention of the people who will have to perform it, creating a loyalty to the Plan as such. One should not underestimate the genuine altruism by which people are motivated in every society ; in modern western society there are a number of points at which it has been thoroughly institutionalized—for instance in the "solidary collectivity...between doctor and patient when the patient is being treated."[26]

Can a society work on this kind of goodwill throughout—without the spur of profit nor the fear of punishment ? That of classical India seemed to do so, and not only persisted for thousands of years but was highly successful. Its economic mainspring was duty. The duty of the Barber was to cut his clients' hair ; of the Potter to make their clay jars and pots ; of the Watchman to stand guard at night. The clients were typically Farmers ; it was their duty to divide the harvest and provide some of it to each of the family servants, not in payment for services rendered but so that all might live. The system was durable ; it had elaborate social means of preventing the ingression of that self-seeking which in modern times has converted services and goods into commodities on the market. It was reinforced by all the sanctions of religion ; the man at the low end of the caste scale was there because of misdemeanours in an earlier existence. He had only one remedy : to do his duty faithfully so that in his next existence the turn of the wheel would be more favourable. Pure reciprocity prevailed,[27] and exchange played a minor role.

The collection of sentiments that characterized ancient India appear often enough today, and attempts are made to combine these with elements of exchange. Among the several Asian leaders it is perhaps U. Nu who appeals to duty and goodwill most explicitly and most copiously.

> For our masses to come up to the standards of living befitting the natural wealth of our country we must all work to the limit from this day and hour. If we all resolve to do so, with one mind, I shall predict that within five years at the longest, this country of Burma will see once again such a golden age as Pagan, in the days when Shin Issa Gawna was exercising the magic of his alchemy.[28]

---

[26] Talcott Parsons and Neil S. Smelser, *Economy and Society*, Free Press, 1956, p. 152.

[27] Karl Polanyi and others, *Trade and Market in the Early Empires*, Free Press, 1960.

[28] From *Towards Peace and Democracy*, quoted in Hugh Tinker, *The Union of Burma*, Oxford University Press, 1957, p. 94.

Unfortunately various elements of self-interest have become institutionalized since the golden age of Pagan, and the individual Burman is distracted from that single-mindedness which may have been possible in earlier times. National plans may be more elaborate than this oratorical proposal of U Nu, but where they do not take account of the institutions and incentives which actually exist in the milieu where they are to be implemented, they may have an effect which is hardly more enduring on their audience than the excitement aroused by such a speech.

The awkwardness of the preceding classification reveals some of the difficulties of an analysis of motivation. I put medals under incentives, suggesting that the Soviet hero of labour would work to get one for himself, just as he would work to get money or a Buick. But from another point of view the medal is simply a token of the fact that the hero has worked hard at doing something which the planners deemed to be in the interest of the whole collectivity—it is a testimony, therefore, to his altruism. And conversely, the profit or the overtime pay that the American puts into a car or house has at least in part the function of announcing his productivity. And what about the Indian who by virtue of his desire for a better place in the next world is moved to carry out a disagreeable but useful task in this one ? Is he motivated egotistically or altruistically ?

Force. Perhaps a clearer separation can be made between either of these and force or coercion. "La contrainte—'ce bien économique oublié," exclaims Francois Perroux.[29] Kenneth E. Boulding suggests a symmetrical relation : "An exchange system is based on a transfer of goods, a threat system on the transfer of 'bads.' "[30] We will not work out here an economics of force, but assume that this can be done. Pressumably one would have to include the threat of force, just as in a consideration of inducement or self-interest systems one includes the promise of a weekly salary which, when it materializes on pay-day, is in the form of money which in turn is only a promise of goods. The promise of goods can be weakened by inflation of the currency ; the threat system can be similarly inflated when those who make the threats do not have the power to make good in a sufficient number of instances. Unconvincing threats, like unredeemable money or vain promises, are issued by a government under strain because its responsibilities exceed its authority ; its constituents expect from it a performance that they will not permit it to attain. It is interesting that the legitimate use of violence is everywhere a monopoly of the state ; in an age when violence is repugnant the state is sufficiently above all other institutions that it stands virtually alone as the repository of this one of our three means. Within varying limits all groups within the state and all individuals are entitled to use the first two of our three modes for influencing individual behaviour ; the state alone is entitled to use all three. It is presumably in part strength derived from this monopoly of force which people have in mind when they assert that because the task of development is so difficult only the state can carry it out.

---

[29] Cited by André Marchal, *La Pensée économique en France depuis* 1945, Presses Universitaires de France, 1953, p. 92.

[30] "The Relations of Economic, Political, and Social Systems," *Social and Economic Studies*, Vol. XI No. 4, December 1962, p. 351.

Aside from all moral considerations, this is the one of our three instruments which is perhaps the most limited in use. In the first place, and to the regret of tyrants in all ages, only so much force can be applied to any individual; after that he dies. But long before that point he ceases to do his best work. As Arnold Toynbee says,[31] it is possible by force to make a galley slave row or a plantation worker hoe, but most of the jobs in industrial society are more skilled, and skilled work cannot be exacted by force. The Soviets made absenteeism a criminal offence in the 1930's,[32] but during recent decades the law has been enforced less and less, and in 1956 violations of labour discipline ceased to be criminal offences.[33]

Nonetheless there is a way in which force is very usefully (from the viewpoint of the planners) added to the complex of motivations pushing and pulling the Soviet manager. The latter needs to make an illegal purchase of materials in order to attain his target; he finds himself in the position where if he is successful all will be forgiven, but if he fails his "crime" will be discovered and he will go to jail.[34] No stronger subjective pressure to succeed could be devised.

In comparison with the psychological ingenuity of this use of force by the Soviets—this is not the place to discuss its morality—the use of force in liberal societies is restricted to certain boundaries of the area of action. *The Economic Development of Malaya* mentions forfeiture of land as a penalty for non-use of it.[35] *India 1962* discusses "ticketless travel" on the railways and its punishment.[36] It is questionable whether these add as much to production as their enforcement deducts.

THE ALIGNMENT OF SELF-INTEREST, DUTY, AND COERCION

In discussing the three elements of motivation separately we have arrived at what seems at first glance the essence of any attempt of the planners to get development under way—the alignment of these impulses. Self-interest, duty, and coercion must all impel the individual in the same direction, and that direction must be material production. In the west this seems to have been associated with changes in religion whose cause is as difficult to ascertain as their effect was clear and beneficial. For Calvinism, the destiny of each one of us is locked in the transcendental, to be finally made known only after our participation in this life is over; in the meantime, however, heaven may give a sign that we are predestined for salvation by causing our worldly affairs to prosper. Fatalism of this variety in no way inhibits striving in the world, but on the contrary encourages it mightily; heaven shows us to be of the elect both by giving us the energy to work hard and by blessing our efforts with success. Though good works themselves create no merit directly in the divine record, they do testify to the sincerity of faith and the gift of grace. Because wealth has no interest for its own sake, intensive striving is accompanied by an asceticism which greatly supports saving and capital accumulation. Evidence of success both brings the admiration of other men and is seen as a service to them ; hence self-interest and duty to others are combined in favour of production. And a further impulse in the same direction is given by threats : since the person who has not succeeded shows thereby that he does not have the favour of heaven, it is permissible for men to treat him severely, or at best to disregard him. This religious impulse towards asceticism and material achievement was found at its height in 16th century Geneva, 17th century Scotland, and 18th century New England, and it provided an essential background for the Industrial Revolution in which all three means of persuasion were applied in the same direction—that of organizing material progress.

---

[31] Cited in Charles R. Walker, *Modern Technology and Civilization*, McGraw-Hill, 1962, p. 347.

[32] Moore, *Soviet Politics*, p. 36.

[33] Jerzy G. Gliksman, "Soviet Labor and the Question of Productivity," *Monthly Labour Review*, Vol. 80, No. 6, June 1957, p. 705.

[34] Reinhard Bendix, "The Cultural and Political Setting of Economic Rationality in Eastern and Western Europe," in Gregory Grossman (Ed.), *Value and Plan*, University of California Press, 1960, pp. 259–260.

[35] International Bank for Reconstruction and Development, Johns Hopkins, 1955, p. 130.

[36] p.350.

The Soviets have achieved the same alignment of pressures on the individual through their purely terrestrial framing of the environment within which his action takes place. How the regime plays on all motives is seen in Joseph S. Berliner's remarks on the Soviet manager :

> The director of a coal mine, for example, whose base salary may be about 10,000 rubles a month, would earn a total of 25,000 rubles in any month in which he overfulfilled his plan by 5 per cent...Personal advancement, public recognition, and all the prestige which the state can officially confer or channel are available to the manager who regularly overfulfils his plans. Demotion is the usual lot of a manager who proves incapable of meeting production targets. ... Soviet managers work hard, not only in terms of their hours of work but in the energy and worry they devote to their jobs.[37]

Furthermore, a kind of ratchet keeps up the pressure—the succeeding plan target is raised if the present one is overfulfilled. Turnover—most of it involuntary—is a measure of the pressure. "In the coal industry, for example, about 40 per cent of the heads and chief engineers of the mines and some 50 per cent of the sector managers change each year." Gliksman also stresses the combination of self-interest, ridicule, force, ideology : "Recognizing that men's impulses never stem from a single source, the Soviet policy has been to apply a combination of all the approaches outlined."[38]

The first part of the thesis of this paper, then, is that planning systems may be discussed and their effectiveness considered in terms of the degree to which the planners are able to ensure that the three sorts of impulse move people in the same direction—that the forces of egotism, altruism, and coercion are aligned. Instances where planning does not work are due to absence of authority, of skill, or of will on the part of the authorities to make concomitant and suitable use of egotism, altruism, and force. In one case the instrument of alignment was religion ; in another it was the encadrement of an entire country, including every factory of any importance, by the Communist Party. Every manager and every worker was aware of being closely watched by other men, if not by God. The likelihood that virtue or negligence in production would be overlooked was minimal. The Communist system is tight enough that the individual is entirely surrounded by it; there are few cracks through which he can make a profit by doing things that the regime does not want done.

---

[37] "A Problem in Soviet Business Administration," pp. 90–91.

[38] Gliksman, *op. cit.*, p. 704.

How different is the condition of an underdeveloped country. The citizen who is a public servant is urged to do his job, but he gets neither material rewards nor genuine respect from his fellows for doing it. Material incentives on the other hand are provided unintentionally for doing things which are hardly contemplated by the national plan—like selling import permits, or other forms of corruption. The public servant comes to have an interest in economic restrictions of every kind so that he can trade on the loopholes in them. The authorities constantly make appeals to his altruism and patriotism; no one can say this is a weak force, but it can be offset by a material interest which pulls in another direction. The spontaneous admiration of a man's neighbours is drawn not by his being a good citizen and obeying the regulations but rather by material success however attained; the demonstration effect determines people's standards for judging their fellows, and not official exhortations. To correct this situation by bringing force to the help of an orientation to the national collectivity—putting an offender in jail—is only possible for a strong and effective government; an underdeveloped government may put a few offenders in jail, but cannot be certain enough of catching the rest to influence general opinion and morality.

The individual who falls under the stress of contrary forces is not peculiar to the underdeveloped countries of Asia. In 17th century Scotland the magistrates fought the high price of wheat by ordering every farmer to bring to market all the surplus grain in his possession beyond what was needed to keep his family until the next harvest, but the contrary force of market rewards for disobeying this injunction made it ineffective.[39] Such a criss-crossing of motives with the resulting check on production is not to be found in either the severely administered totalitarian economy or the liberal economy; it is a special hazard of the emerging country—European or Asian—which is not yet either capitalist or socialist.

Thus the environment of the individual pushes him in various directions at the same time. The exhortations of his government play on his altruism, i.e. his sense of the whole national community; the temptations of black-market profits play on his loyalty to his family, to his particular ethnic group or other collectivity smaller than the nation. The threats by which a government which is relatively weak tries to coerce him in the interest of the nation and away from the interest of any smaller collectivity lack the high probability of execution, the severity, and the grounding in the mores on which effective coercion depends. But now we must pass from the forces acting on the individual to their homologue, as Parsons calls it, in the community. Individual motivations arise in social activities.

---

[39] Donald Grove Barnes, *A History of the English Corn Laws from 1660–1846*, Reprinted, New York: August M. Kelley, 1961, p. 33.

INDIVIDUAL MOTIVES ARE RELATED TO SOCIAL STRUCTURES

We have spoken as though planning was a matter of the manipulation of individuals. In fact this is a very incomplete — as well as morally repulsive — perspective ; the actions of individuals within groups are responsible for the meanings which take shape in their minds and without which they cannot be influenced through the means we have mentioned, or through any other means. A medal awarded by the state will be an incentive in the degree in which it is interpreted as such by the recipient and his associates. Their praise — or mockery — will gear into ego's motivation system. What is true of medals is true of money and goods — they must be interpreted as symbols of the gratitude of the collectivity for good work done on behalf of all. If people can only be induced to work by the intrinsic material value of the goods they are offered, then development will never take place — it will be too expensive, for there are not enough such goods available in its early stages. In the study of the three sorts of individual motives as they act in practice, it is necessary to look at what we may call the social environment, a terrain which is not geographical but organizational, of the underdeveloped country.

The social terrain becomes increasingly differentiated. The social terrain of the presently undeveloped countries is already richly marked both by tradition and modern forms of grouping. Consider the people living in an Asian village, a territorial collectivity important for a number of purposes. In addition a group of villages may be related to one another in having a common landlord with various sorts of property and non-property relations to his tenants. As a minimum he will collect the taxes they owe, along with a commission for himself ; he may dispense justice in regard to certain matters ; he may provide tools, seed, loans, and food. Aside from this the peasants may be formed by caste regulations into conscious homogeneous occupational groups. They are likely in addition to be members of a religious community, and will certainly be oriented to smaller and larger kinship groups. To describe these organizations in detail, along with the sets of ideas associated with them, would be to describe a peasant society.

But to talk seriously about economic development one cannot stand as far back in history as we have done in the preceding paragraph. The classical peasant society has everywhere been altered by colonial contact, as well as by post-colonial social change. The groupings which made up unspoiled peasant society all survive in one condition of repair or another, but today they exist alongside of modern forms. New syncretistic religious bodies ; political parties ; political protest movements that have not yet become parties ; the following of an important individual in a village or town, which acts in common defence and can be called a faction ;[40] trade unions ; members of a housing block ; producer, consumer, and credit cooperatives ; government agencies ; military detachments which engage in a continuing task over a long period ; schools and universities ; all these and many others are today superposed on the original peasant configurations. The newer groupings are distinguished from the older in permitting more individual mobility ; they encourage a person to consider where he stands in a number of social hierarchies, and they incite him to try to climb to a higher level ; he sees how his progress in one of the hierarchies may help him in the climb through another. A vast pressure of upward mobility is created by these modern structures, especially in the large cities of developing countries. The personal ambition which is stimulated in its turn promotes further structuring ; a man whose ambition has been aroused by his attempt to rise in a political party, and then frustrated by insufficient opportunity within it, will go out and found his own party. It is by this process that an ambitious urban elite is created in all the great cities of the world, an elite which everywhere provides the chief push for economic development.

The groups block each other. Now this congeries of groups pull in different directions in such fashion as to prevent each other from moving. The underdeveloped country looks to an outsider as though it is lacking in dynamism. This is a false impression ; its citizens are ambitious ; they are working hard as individuals and in association ; if they have fewer professional organizations they may well have more political ones than the advanced country. But the organizations block one another.

As an example consider what we may call a constellation of interests having an area reference, those groups and persons having a common stake in a certain locality. The interests clustered about the harbour of Colombo in Ceylon will serve as an example ; they include the stevedores who load and unload cargoes, and the managers and owners of the stevedoring agencies ; the harbour management itself, an agency of the state but not thereby slavishly identifying its interest with that of the state ; the owners of warehouses in the neighbourhood of the harbour ; other real estate interests in the city of Colombo ; its news-papers, city council, etc. We can go a considerable distance in ever widening circles without ever fully taking in all the interests that belong to this constellation. It is enough here to note that they are extensive but much smaller than the nation of Ceylon.

Now from the widest viewpoint it is highly desirable that the nation not be dependent on a single harbour. Quite aside from the danger of an emergency, day-today economy dictates that there be a number of points of entry into the country. The report of the International Bank for Reconstruction and Development, undoubtedly the most thorough single study of the economy of the island ever published between two covers, is emphatic on this, and proposes Galle as a second harbour. "For both strategic and economic reasons the long-term development of Galle harbour into a modern port providing safe anchorage in all weathers ought, we believe, to be continuously borne in mind."[41] The reason for this recommendation is that "Galle is the principal town of the highly productive Southern Province. Because of the [present] limitations of the port, most of the region's imports and exports are routed through Colombo, 75 miles away. If any appreciable volume of trade were to be diverted to Galle there would be a corresponding decrease through Colombo."

---

[40] Oscar Lewis, *Village Life in Northern India*, University of Illinois Press, 1958, Chapter 4, pp. 113 -154.

This diversion, a merit to the IBRD team which is studying Ceylon as a whole and observes the overcrowding in Colombo and consequent delays to shipping, is a demerit for the constellation of interests to which I have referred. In this regard the situation is similar to that in the developed world; New York did not want a seaway that would bring ocean traffic to Chicago. The difference between the developed society and the underdeveloped one is that in the latter the institutional channels for the free assertion of new interests have not yet been created; such new interests as arise are weak and readily blocked by existing interests. Development consists in the arrangement of means by which deadlocks can be broken to the advantage of new organizations.

The groups in the interior of the national state of which we are speaking include some who by their very nature oppose the goals of the planners. Landlords exercising various sorts of feudal authority are the most prominent of the traditional forces, and much of the energy of the urban elite is devoted to uprooting them. The interest of the landlord is the opposite of that of the urban elite; he wants a weak government so that he can sell his surplus crops abroad and use the proceeds for such luxury consumption as cars and foreign travel. To build local industry which will ultimately produce cheap goods for mass consumption interests him not at all. In the measure in which ancient modes of cooperation continue, the landlord has his peasants on his side; even if they are sophisticated enough to reject his exploitation they are not sure how much better off they will be under a development-minded elite whose promise to raise their income is far in the future. In the peasant perspective this promise is hidden by the much more imminent austerity of the saving and investment which are a pre-condition for development. Similar analyses may be made for privileged castes and other traditional groups which see themselves as likely to suffer at least short-term material losses through energetic development, and who even in the long run would lose in relative position because of the more rapid advance of those presently below them. But there is no need to elaborate on this theme since everyone is aware of it.

It is less apparent that modern sorts of organization can constitute roadblocks of a similar kind. The matter is complicated to express because any given type of organization can spur or defeat development. We spoke of the constellation of interests around Colombo harbour as having enough power in the national society as a whole to defeat the development of a rival harbour; that this is possible is due to the centralized fiscal system. On the other hand, the six major ports of India engage in active competition for traffic, despite the fact that they are all under the control of the national government and three of them are run directly by the Central Government. One assumes that the attractions of monopoly are not lost on any of them; all would presumably like to block the others; evidently none is in a position to do so.[42] This duality of modern types of organization is seen clearly in trade unions: on the one hand these are nearly indispensable means of securing the recruitment and commitment of labour; on the other they can advance the cause of labour to the point where management is defeated in its attempt to produce. It would be instructive to compare in detail the means used to overcome this difficulty in the United States and the USSR.

---

[41] *The Economic Development of Ceylon*, Johns Hopkins, 1953 p. 624.

And what about the struggle between a central government and particular ethnic groups—that of Indonesia and its Chinese citizens, Ceylon and its Tamils, or India and its Marwaris, Nagas, Nayars, etc? The situation varies greatly among these cases, the only common element being the obstacle to the desired unity of the state presented by the culturally differentiated elements. The members of these have a consciousness of themselves as a group, by which the loyalty they bear to the state is in greater or less degree attenuated. In relation to the state this loyalty to a sub-group is called selfishness. With more or less insistence the modernizing elite have taken the view that the state can tolerate no rivals, and have tried to block any that appeared. Presumably a better policy would be to admit the several sub-national elements to open competition and arrange that the focus of competition be production. If the Chinese of Indonesia are getting rich on usury, block that by all means, but do not stop them from becoming rich on manufacturing. Since for the time being they have little identification with the state, it is even less possible to call on their altruism than on that of the members of the majority culture; one must harness their effort by the attraction of self-interest. At least one Indian scholar, M. N. Srinivas, sees a hierarchy of groups and of values. "A man stands for his village in relation to other neighbouring villages, his taluka in relation to other talukas, and so on...He is also an Indian in relation to non-Indians...Tensions and conflicts maintain the identity and separateness of groups of the same order but these groups can and do unite at a higher level."[43]

THE DEVICE THAT KEEPS THE SOVIET BUREAUCRACY DYNAMIC

The Soviets have been successful in putting inter-organization rivalry to work, and especially the rivalry between the hierarchical organization of the state and that of the Communist Party. The State works down from the legislature with its Presidium or executive committee; the ministries of production (in part replaced since 1957 by regional councils or *sovnarkhozy*); and so on to the factory and the individual division of the factory, in an ordered chain of command. The lower reaches of any such chain of command can become frozen in the observance of rules to the neglect of production, and be effectively beyond the authority and stimulation of the higher reaches. In a bureaucratic system progress tends to be neglected to the benefit of stability. But in fact the whole Soviet state apparatus is interwoven with, paralleled by, and infiltrated through the Communist Party. Some 9,000,000 strategic individuals who are Party members report upwards through the separate Party chain of command which has its own loyalties. Stalin attained supreme power when he became both Prime Minister of the State and Secretary of the Party, so joining in himself the two hierarchies. But though they overlap they are not entirely amalgamated at any lower level. Party members are present in all of the 200,000 factories and 100,000 construction sites of the Soviet Union, though not necessarily in formal charge. They ensure that no manager's devotion, initiative, or delinquency goes unobserved. Their job is to exert constant pressure towards production, as well as to ensure Party political control. Only through very elaborate collusion could progress be halted, and though this occurs, there are enough eager and ambitious subordinate members of both hierarchies that the collusion is likely to be denounced sooner rather than later.

---

[42] *India* 1962, p. 358.

[43] "The Nature of the Problem of Indian Unity," *The Economic Weekly*, Bombay, April 26, 1958. Reprinted in Myron Weiner (Ed.), *The Civilization of India*, Vol. I, University of Chicago Press, 1961, p. 53.

I repeat that this essay does not contain the answer to the problem of development; at most it contains a statement of the question. There are nonetheless two general aspects of development which ought to guide all who search for an answer: the disembedding of the economy, and the personality of the enterprise.

DISEMBEDDING THE ECONOMY IS A PRECONDITION FOR RATIONALITY

In most times and places, "man's economy...is submerged in his social relationships."[44] The arrangement described above for the village of classical India is probably more typical of the main historical types than any model based on exchange. "The characteristic feature of primitive economies is the absence of any desire to make profits from production or exchange."[45] And as long as men grow their food while they carry on family life and religious observance, as long as they enjoy relaxation, association, and play when they plant and harvest, there is no possibility of applying rationality, for the economic portion of the activity is not a separate entity; there is no case for changing the methods of planting merely to double the crop, since the crop is only a part of the purpose of the activity. Will the interests of the family, of religion, and of play be equally satisfied by a new method? This is never easy to determine. Parsons and Smelser put the matter more moderately, but their intention is clear enough: in eastern peasant societies "the typical organizational unit...has been a *combination* household and productive unit in which the differentiation between occupational and familiar roles does not hold. Economic considerations are certainly relevant to the analysis of behaviour in such a unit, but inevitably it is subject to constraints on giving primacy to economic considerations, constraints which the modern Western firm does not face."[46]

---

[44] Karl Polanyi, *The Great Transformation*, Beacon Press, 1944, p. 46.

[45] R. Thurnwald, *Economics in Primitive Communities*, 1932, p. xiii (cited in Polanyi, *Great Transformation*).

[46] *Economy and Society*, p. 81.

On this description village economic activity has no costs and brings no profits. Accounting considerations are irrelevant where there are no choices to be made of which calculated cost could be the criterion. Alternative uses of resources are not in question ; hence there is no scarcity. It seems paradoxical that the more primitive the society the less scarce are its factors of production, but this is true nonetheless. Scarcity and cost come in general with increasing wealth.

One cannot even think seriously about incentives to individuals while they are engaged in a satisfactory traditional division of labour with their kin and neighbours. Development includes drawing their attention to the material elements of their social life as such ; it involves some measure of creation of (symbolic and other) satisfactions derivable from greater quantities of these, persuading people to set aside a portion of their day for "work" within which the cost calculus will apply, and allowing all other time to be regarded as devoted to play, hobbies, religion, etc. and excluded from the cost calculus, whether or not it may produce something incidentally—this is to pursue the disembedding of the economy. As far as anyone can tell, this disembedding is the necessary condition for mastery of the physical world in a modern, flexible, changing division of labour.

That the disembedding is never complete appears with consideration of such facts as that the American corporation—perhaps the closest approximation to a purely economic institution yet developed—still is concerned with the health of its employees, still provides amenities to its senior employees on expense accounts, decorates its grounds and buildings, all beyond the strict requirements of production. James Abegglen finds that the affiliation between employee and factory in Japan is an enduring one, to the disregard of possible alternative employment for the former in good times and of the financial interest to the latter in dismissing him in bad times.[47] We do not here take up the classification of disembedding, nor the interesting question of the minimum degree that would permit choice and hence rationality in production. In a country in which the process of disembedding has not started, the planners could with difficulty initiate it—indeed they would be helpless. But luckily the planners in most countries today have instances of going industries within their borders ; their job is to shift resources towards these, so that they can stand out from the confused social landscape in which they now struggle, and expand and be imitated. We elaborate this as the last point of our argument.

DEVELOPMENT IS THE PREDOMINANCE OF ORGANIZATIONS GEARED TO PRODUCTION

We have pictured the social terrain of the developing country (like that of the developed one) as made up of a host of organizations : religious, political, social, economic. These severally focus the loyalties of their members. Within each of them the desire for praise and money, as well as the sense of duty, arouses ambitious individuals to compete for eminence and authority. Between them a state of war often exists. This rivalry can be useful, on the all-important condition that none be in a position to block the progress of others. The problem is to strengthen such of these organizations as are productive, to enable them to expand at the expense of less efficient rivals. In the degree in which the economy becomes disembedded this is easier and easier to arrange; the productive enterprise arises which has renounced all interest in religion, and religious organizations lose their interest in production. In the limit, in Parsons' phrase, each organization comes to operate more and more within its own sub-system of the social system, and it is then possible for the planner to turn his attention to economic organizations without interfering with other aspects of the national life. But the attention of the planner must be restrained or it will do more harm than good.

---

[47] *The Japanese Factory*, Massachusetts Institute of Technology, 1959.

The need for freedom of the production unit is well expressed by J. K. Galbraith when he speaks of the corporate personality of the unit.[48] It is inevitable, he says, that modern production will be carried out by organizations similar to those of the United States or the USSR, essentially autonomous entities, developing each its own character in the course of struggle with the necessarily difficult technical and organizational problems of production. "Organizations develop personalities that are somewhat similar to the personalities of individuals. It is not an accident that the law regards a corporation as a fictitious person," says Kenneth Boulding. "Thus, organizations have constitutions, rules, procedures, precedents, organization charts, balance sheets, and formal structures that are analogous to the body of a natural person."[49]

The corporate matrix is fundamental to development because it is within the corporation that common activities generate those responses to the stimuli of self-interest, duty, and coercion of which we have spoken. The corporation, more than most other groups outside the family, forms the personalities of its participants; it attaches to definite objects the initial vague tendencies to self-interest. That the person seems "instinctively" motivated to seek money, a title, a distinctive uniform, a carpet on the floor of his office, a seat at the controlling end rather than one of the receiving ends of the office buzzer system, arises in what may be called secondary socialization. The sort of person one becomes, the sorts of wishes that make up one's personality, are in considerable part determined by the activities within the organization in which one spends his working life.

The main stream of economic thought has taken account of this, and different writers have tended to concentrate on one aspect or another of the corporation or firm. The entrepreneur who brings the elements together is recognized as far back as J. B. Say's writing. Alfred Marshall develops further some of the qualities of the entrepreneur in his discussion of business management.[50] Schumpeter is perhaps the best-known exponent of the view of the entrepreneur as the dynamic agent of production, who breaks a circular flow and makes his one-time contribution, for which he receives a one-time profit. As entrepreneur he has neither capital nor skill in mechanical inventions ; he works more by will than by intellect ; he buys people rather than persuades them[51] ; the means to buy them he secures as credit and "rides to success on his debts."[52] He is not of the established order : "in general it is not the owner of stage-coaches who builds railways."[53]

---

[48] *Economic Development in Perspective*, Harvard University Press, 1962, Chapter V, "The Instrument of Production," pp. 60–74.

[49] *Conflict and Defence*, Harper and Bros., 1962, p. 178.

[50] *Economics of Industry*, Macmillan, 1928, pp. 162–170.

Our view is much less heroic. Undoubtedly the leadership and force of character of individual man has had an important effect on industrial expansion, and Schumpeter was writing in an era of genuine enough captains of industry. But since then a managerial revolution has occurred ; the owners of business have been pushed even further from its control,[54] but management and innovation are institutionalized. While the modern corporation needs all the entrepreneurial talent it can get, it must also contrive to use less original minds in its complex procedures of decision and coordination. Under the collective impulse to innovation, once the right atmosphere has been created sound and original decisions will be made by groups of men who individually would not have attracted Schumpeter's admiration.

The firm is thus by no means simply its manager writ large, for a good manager does not try to make all the decisions ; decision-making is inherent in the whole corporate being. Progress is being made in the study of decision-making within organizations,[55] and some principles are beginning to appear.

If the manager does not run the business, in the sense of telling everyone what to do, even less can an outside planner, ministry, parliament, or group of shareholders run it. Attempts to influence its internal life from the outside—well-and ill-intentioned are on the same footing here—will damage its coordination, its capacity to make those quick and relatively sound decisions that are the essence of high productivity.

But there can be no question of extending the freedom of the corporation to a license to produce or not at its whim. In fact the whole pressure of the outside society must be directed to insisting that it produce. Those of proven capacity to produce secure society's support and grow larger; those without this ability must be deprived of support so that they will diminish. Whether the goal is expressed in a balance sheet or in the volume of physical production is secondary to the importance of imposing on the concern the obligation of meeting a relatively objective test of efficiency. The operative point is that the planners set objectives in the simplest possible terms, allocate amounts of capital, and then simply disappear from the picture. At the end of each budget period one can see how far the skill of the corporate body, its collective energy and coordination, have succeeded in overfulfilling its objective or earning more than the market rate on its capital. Aside from prizes and every sort of symbolic encouragement, the planners will see that a steady stream of the precious capital at their disposal goes to the entity which succeeds.

---

[51] *The Theory of Economic Development*, Oxford University Press, 1961, pp. 88, 89.

[52] *Ibid.*, p. 70.

[53] *Ibid.*, p. 66.

[54] Adolf A. Berle, *Power Without Property*, Harcourt, Brace, 1959.

[55] R. Likert, in Mason Haire (Ed.), *Modern Organization Theory*, John Wiley and Sons, 1959, pp. 196–197.

Like Galbraith, I have tried to express myself in the preceding paragraphs in a way which would apply indifferently to any economic system. The main economic difference between the people's capitalism of the United States and Soviet communism is that in the former the capital which is invested is made up principally of the undistributed profits and depreciation allowances of several thousand large corporations, along with what they can borrow from banks and other sources, while in the latter most of the capital—70% according to Berle—has been removed by the State planners from the concerns that earned it and reallocated according to the planners' view of where it would best be used. Control of capital is power, and the protection of the citizen against the abuse of power is possible in only one way—by dividing it up among many holders—but this non-economic consideration is not our main interest here. In both systems matters concerned with production are dictated by the exigencies of survival of the organization. "Resistance of a small bureau [in the USSR] against absorption into a larger one is motivated almost exactly as is the struggle of a small American corporation to maintain its share of the market against some other giant who proposes to expand. Competition within the Soviet system is cast in the bureaucratic rather than in the commercial mould. Yet the stakes are much the same."[56] In both systems there are high income and prestige for the successful manager, and limited income or prestige available to individuals outside the productive system. In both there is an accelerating tendency for the owners—stockholders in the United States, the State in the USSR—to have less and less influence on the concern's internal affairs as long as output exceeds input.

STRONG ORGANIZATIONS ARE IMPORTANT IN ALL FIELDS

The task of the planners is to find productive organizations in all fields, and to provide them with capital, to water with the flow of resources the sorts of organizations that have achieved an internal arrangement permitting effective cooperation, that can adapt to the ever-changing exigencies of production. No planner as such can create such an efficient organization. The mode of operating which we call efficient, its dynamism, are as much beyond the conscious, rational instrumentalities of a central planner as is any other natural social organism. Any attempt even to improve existing organizations by imposing detailed rules on them will only do harm. The planner is helpless in any country in which there are not at least a few effective organizations; such countries have not yet started to develop. But once there are even a few, the task of the planner comes within human capabilities—it is selective support, as far as possible according to objective criteria such as production and profit. The businesslike character of the organization will imprint itself on the social being of its participants; industrial man is created in efficient industrial organizations.

---

[56] Berle, op. cit., p. 144.

Francois Perroux recognizes this when he says, "Former des hommes économiquement efficients reste une tâche urgente des sociétés viellies et devient la tâche par excellence des sociétés naissantes,"[57] but he gives less attention to the degree in which such men are formed by cooperation in purposeful activity. The commonplace notion that efficient individuals make efficient enterprises contains an element of truth, but there is more truth in the proposition that efficient enterprises generate efficient men. The internal pressures in an efficient enterprise act on the people who participate in it, especially its younger members, in such fashion as to create the attitudes and capacities which it requires. The process by which the young animal which is born into a family is made into a human child is called socialization; it includes the learning of language and other elements of the general culture. The process by which the human adolescent is made into a member of the work force with technical skills, a professional conscience, and the host of other attributes that constitute him an effective member of a team, may be called the secondary socialization of industrial man. This paper suggests that the principal means to it is not exhortation, nor even schooling, but engaging in productive activity.

But the criterion of production would not seem to apply to such institutions as government departments, universities or research agencies. Their efficiency must be judged by means which are frankly subjective; output of students will tell as little about them as will number of patents issued, since quality in students as in patents is unspecifiable in objective terms. The use of judgment is inescapable, and it must focus on whether the activities in which the organization engages are of so purposeful a character that they result in an "output" of valuable research, etc. on the one hand, and suitably form their participants on the other. Aside from this they must have complete autonomy. All the remaining conditions we have set down apply to these just as they do to organizations in the most standardized of production. The universities etc. which are given more capital will form more personnel in the image of the personnel already there; their example will influence other organizations to resemble them. Hence the need for those whose spread is desirable to be those which secure capital.

CONCLUSION

We started with the planners—individuals associated with the urban elite of the underdeveloped countries—and discussed briefly their preferences, as well as the elements of social power—capital in a broad sense—which they dispose for giving weight to these preferences within their societies. The instruments of motivation through which the plan can influence individuals seem to be classifiable as self-interest, duty, and coercion. But attempts to influence people directly constitute a kind of manipulation not likely to be successful. The individuals must be seen as factors in social groups or organizations, some more and some less productive, some not concerned with production at all. The main preoccupation of the planner, aside from the technical consistency of his plan which we take for granted, must be the allocation of capital (or social power) to those organizations that have shown by past success their capacity to convert a given input into maximum output. As these effective teams gradually rise in the social landscape above inefficient producers and organizations devoted to other purposes than production, such as caste, the activities within them will modify the attitudes and skills of their human components, and more organizations will adopt their form. This is the process of development, whose exigencies cannot be escaped by refuge in any ideology so far invented.

---

[57] *L'Economie des jeunes nations*, Presses Universitaires de France, 1962, p. 201.

THE THREE BASIC PURPOSES OF AN ECONOMIC SYSTEM

- **To achieve optimum growth**
- **To apportion production wisely**
- **To maintain economic justice**

By LEON H. KEYSERLING

Former Chairman, Council of Economic Advisers to the President; Consulting Economist and Attorney; President, Conference on Economic Progress U.S.A.

INTRODUCTION

During the period that I have known Professor Mahalanobis, both in India and in the United States, I have developed an ever-increasing appreciation of his intellect, unusually wide range of interests and abilities, and devotion to the progress of his country. One of the outstanding and most appealing attributes of Professor Mahalanobis is his deep concern about, and his desire to enlarge his knowledge and understanding of, what is happening in all of the countries of the world. This is not only because he believes that every variety of current human experience may have something to offer to India in the course of its economic advance under a system of freedom, but also because he values the whole of human experience for its own sake. He is thus both nationalist and cosmopolitan. For these reasons, I have thought that I could best pay tribute to him on his Seventieth Birthday by offering for inclusion in the volume commemorating that event some comments which I have made about the economic problems of the United States—some parts of which may be relevant elsewhere—before the most important economic committee of the Congress of the United States.* I would like to add also my warm good wishes to Professor Mahalanobis, coupled with the hope that he will have many more years of good health and high usefulness.

BASIC REAPPRAISAL OF ECONOMIC THOUGHT AND ACTION

The basic reappraisal of American economic thought and action at mid-century now being undertaken by the Joint Economic Committee has not only domestic but worldwide significance. While this reappraisal has been called for most insistently since the launching of the first sputnik, it has been needed for a very long time. The advent of the great depression caught us unprepared. The changes in economic thought and action during the period of recovery therefrom were useful but not sufficient. They produced only an incomplete recovery prior to World War II, and in any event they are not entirely relevant to the problems of midcentury. The economic thought and action applied during World War II, and to an extent during the Korean war, have been largely sloughed off since then. While to an extent this has been appropriate, the lessons of wartime experience are not being applied sufficiently today. Thus we now find ourselves, despite the presence of problems as critical and demanding as any the nation has ever faced, with a dangerous lack of systematic economic thought and action attuned to the times.

---

*The Congressional Joint Economic Committee, established under the Employment Act of 1946.

We are still indulging in classical economic approaches which never correctly described nor adequately dealt with the problems of the periods in which these approaches were originally developed, plus economic approaches which were once valid but are now outdated, plus pseudoeconomic approaches which are mere shibboleths or slogans to incite emotions or to take political quick tricks rather than to achieve effective action. The people and the nation have paid dearly, and are likely to pay even more dearly, if this indulgence continues.

I shall divide my presentation as follows: (1) an attempted restatement of basic economics for the midcentury; (2) an analysis of the main respects in which we have departed in thought and action from this basic economics.

A RESTATEMENT

It seems to me that the three great purposes of practically any economic system, including ours, at the midcentury are these:

To achieve steady and optimum economic growth in real terms. This means calling forth the maximum use of our productive capabilities, including manpower and brains, technology and science, and natural resources. This comports with my interpretation of the maximum production and maximum employment objectives of the Employment Act of 1946. The whole history of economic progress has been founded upon the use of improved technology to increase production, and especially to increase production per capita. The faster a nation does this, the stronger it becomes in an economic sense, and the better able to do what it needs to do, unless it is burning itself out by going so fast that it depletes its resources. We are in no foreseeable danger of following the latter course;

To apportion our total national production wisely, in accord with relative priorities of needs, so that we do not get what we need least at the expense of what we need most. While the utilization of our total resources is now so slack that to talk about a choice between more consumer gadgets and more of the things we need most is a fundamental misstatement of our current problem, nonetheless we must always make sure that we do not get more gimcrack gadgets and less defense, international economic cooperation, and schools;

To combine economic progress and efficient use of resources with economic justice, as the American people understand economic justice. To illustrate, the first two great purposes might be attained, even while the farm population lived by eating the economic crumbs.

I am prepared to challenge any economist to indicate how we could fall short in any of our valid economic purposes, if these three great purposes are adequately achieved. For example, any particular inflationary trend could be challenged only on the ground that it interfered with or threatened one of these three great purposes. Further, these three great purposes interpenetrate. For example, it is unlikely, within the contours of the American economy, that we can meet the priorities of our national needs without an optimum rate of economic growth.* It is also unlikely that we can maintain an optimum rate of economic growth without meeting these great priorities of our national needs, or without maintaining a distributive system which accords with our concepts of economic justice.

---

*Professor Galbraith's "The Affluent Society" notwithstanding.

The basic conditions. The basic conditions for achieving these three great economic purposes may be stated fairly simply. I do not think that the desire or initiative or enterprise is lacking among our people and their various privately organized efforts to move consistently toward these three great purposes. Certainly, our public institutional devices and our system of free government are equal to their share of the task. However, the central economic condition for the accomplishment of these three great purposes, by no means yet achieved, is that there be a balanced development of (a) investment in the means of expanding production and (b) the private and public consumption—meaning public demand for goods and services—at balanced growth rates which call forth maximum use of total resources and optimum economic growth.

I stress this, because it is quite possible to have an equilibrium between investment and consumption at levels far below maximum use of resources and optimum economic growth. Chronic excessive unemployment is quite conceivable, and in fact we have been suffering this for a number of years and are now in danger of its persistence for a long time ahead. Further essential conditions for the achievement of these three great economic purposes are that the distribution of investment be well-balanced; that the distribution between private and public consumption be consistent with meeting sufficiently, but not excessively, the great priorities which depend upon public programs; and that the distribution of both private consumption and public services be consistent with the advancement of economic justice.

Maximum purchasing power. These salutary economic conditions depend mainly upon maximum purchasing power, as I interpret the true meaning of that term in the Employment Act of 1946. The division of purchasing power between spending and saving is merely a facet of this problem, since saving at one point usually needs to flow into spending at some other point to prevent unemployment. Maximum purchasing power is therefore a volume and distribution of effective purchasing power which brings the various types of private and public spending into line with an investment and consumption pattern consistent with the achievement of the three great purposes of our national economic life as I have set them forth initially.

This flow and distribution of purchasing power is affected by all basic national economic policies, including public spending and taxation, monetary policy, and various other programs including some regulatory programs. Other public programs, including for example the social security system and the wage and hour legislation, importantly affect the private flow of purchasing power. In the field of private economic activity, prices and wages and profits, and in fact practically all conscious private economic policies or automatic private adjustments, are primarily significant in terms of their bearing upon income flow and distribution as these in turn affect the three great purposes of our national economic life. It should be added that public and private developments interact—a fact insufficiently dealt with by current economics. Public policies influence private economic decisions. And private economic decisions can to a large extent "decompensate" for or cancel out changes in tax policy or other public policies.

OUR GROSS DEPARTURE FROM BASIC ECONOMICS

Except in wartime, it is my view that most of our economic difficulties and shortcomings have resulted preponderantly because our economic thought and action have departed seriously from the basic economics outlined above. This point is so important that I shall now set forth what seems to me the outstanding departures, which have done and are still doing so much damage.

(1) *We have thoroughly confused ends and means, which results in a sacrifice of ends, and also results in the choice of the wrong means.* We have not set forth clearly and quantitatively, on both a short-range and a long-range basis, the three great purposes or ends of our economic life. Consequently, we have had no effective way of attuning means initially to the attainment of these ends, nor of readjusting these means pragmatically and quickly when we see that they are improperly attuned to these ends or when changing domestic or world circumstances require that we revaluate the relative priorities of the ends themselves. Instead, in public spending policy, tax policy, monetary policy, and other policies, we frequently apply criteria of judgment in a vacuum unrelated to the ends.

We tend to talk about budgetary surpluses or deficits as good or evil in themselves, when they are nothing of the sort.

We debate endlessly about the relative trends in wages and prices and profits, without examining what relative trends are conducive to the three great purposes of our national economic life, and without recognition that the same relationships among wages and prices and profits—either in terms of absolutes or of trends—are not desirable at all times under all circumstances.

We think that we have proved a lot when we show that the upward movement of prices in recent years has been in the administered price area, without considering in general whether an administered price system on the average serves our economy better than a price system which is competitive to the point of irrationality, and without considering specifically just how an administered series of price increases have impacted upon our great economic purposes or whether a different set of developments would have worked better.

We assume that stability of the general price level is most conducive to the fulfilment of our three great economic purposes, or we even assume that general stability of the price level should be regarded as a prime economic objective on a parity with all others. While I share the view that general stability of the price level is desirable and reasonably attainable, I believe that bad price results and bad economic results both ensue from putting price stability on an absolute parity with all other economic objectives. Giving price stability more weight than other considerations would be disastrous. A careful examination of our economic experience, over many years, seems to indicate clearly that imbalances in the price and income structure are far more serious a problem than the absolute levels or trends of prices *per se*. It is noteworthy, in this connection, that by far the greatest economic calamity in our history resulted from the 1922-29 distortions in the income and investment-consumption structure despite a remarkably stable price level, except for falling farm prices.

(2) *We have not developed and put to use a systematic and integrated portrayal of the capabilities and needs of our economy as a whole, and measured both ends and means in terms of this complete tableau.* I am firmly convinced that to do this was the original intent of the Employment Act of 1946, but the departure from this intent has become very serious in the administration of this act and in the Economic Reports of the President thereunder. But this damaging deficiency is not chargeable only to the government. With a few notable exceptions, American economists in general, as this committee from long experience must fully recognize, have become segmental specialists who try to study more and more about less and less, instead of helping us to obtain a comprehensive view of ends and means. The shibboleths and slogans which affect economics are so prevalent among laymen because they have become so prevalent among the professionals.

The building of so-called economic models has been a partial step in the right direction, but it is only in its initial stages. The models are usually used to depict where we ought to go in future. But they are seldom used to make a discerning analysis of the past, by examining carefully what went wrong, and where actual performance departed from the model. Possibly one of the reasons why many economists have been reluctant to do this is that it would force them to call a spade a spade, and to indicate certain misbehaviours of private and public economic policies which would involve them in controversy or cause them to be called "politicians." Until we get down to this kind of analysis of past developments, the models do not serve one of their main purposes.

Insofar as these models are used to look toward the future rather than the past, they are not usually equated sufficiently with quantitative policy recommendations to tell us just what kind of private and public means will carry us forward toward the glowing objectives. Another difficulty with the models looking toward the future is that they are generally too aggregative in their approach. They tend to assume that, if we need to achieve a certain gross national product, a deficiency in consumption can be made up for by an expansion of private investment or public outlays, or vice versa. This is quite superficial, because it neglects the problem of workable equilibrium, and also tends to neglect the great priorities of our national needs.

(3) *In fiscal policy, our thought and action are among the most serious departures from basic economics as I have defined it.* The current propaganda that Federal spending is *per se* odious, that any and all downward trends in Federal spending are desirable *per se*, and that private spending for more luxurious resort hotels on Miami Beach is

per se preferable to public spending for schools or river developments, is hardly worthy of serious comment.

But even what is now called the modern or advanced view of fiscal policy has many superficial and therefore harmful aspects. The customary formulation that we should run a budget surplus when productive resources are overstrained, and run a deficit when there is excessive economic slack, is a risky generalization when not subjected to further refinement.

For example, there is a tendency toward the view that, if a change in the net budgetary position is desirable, it makes little difference—aside from practical political considerations—whether this change is accomplished through a change in spending rates or a change in tax rates. One defect in this position is the assumption that a change in spending or tax levies of the same size will automatically produce the same (allowing for the multiplier effects) change in the net budgetary position, without regard for the diverse impact of spending and taxation upon the general economy. Another major defect in this position, in my view, is that it disregards the second of the three great purposes of our economic life, namely, the allocation of a sufficient quantity of our total productive resources to the great priorities which require public spending programs. Such a position, indeed, confuses and misstates the respective purposes of public spending and taxation.

Spending. The real purpose of Federal spending is not primarily to stimulate or repress the economy, nor to achieve a budget balance. The towering central purpose of Federal spending is to allocate a sufficient proportion of what our total national output would be, under conditions of maximum employment and production and optimum economic growth, to the great priorities of our national security and domestic public needs which can be served only by Federal spending, or served better by federal spending than in some other way. This means that the level of Federal spending should be determined by estimating first the total potential output of our economy under optimum conditions, and then deciding as a matter of national policy what part of this output we want to allocate through Federal spending to various national priority purposes.

Of course, this allocation should take account of the impact of Federal spending upon the three great purposes of our economic life, and not upon just one of them. It should also take account of our traditional concepts of the division between private and public efforts under our kind of economic system, although we should not freeze this concept to the point where it jeopardizes our national security or starves essential domestic programs. We were realistic about this in wartime, and we should always be realistic about it, even though the results arrived at would be very different in other times than in wartime.

If the national economy as a whole is performing far below optimum levels and has large slack resources, this is certainly no reason for reducing or holding Federal spending to levels below those arrived at by the methods suggested above. It is preposterous beyond words to sacrifice national priorities when idle resources are available to produce them, just because a slack economy is not yielding enough tax revenues to cover all of the bill. Moreover, this preposterous procedure compounds the general economic difficulties of a slack economy instead of reducing these difficulties, and thus is bad even for the net position of the Federal budget in the long run.

On the other hand, if our economy is functioning above optimum use of resources, with excessive strain upon productive capabilities, the level of Federal spending should not be reduced or held below the levels arrived at by the methods suggested above. Any such procedure would be tantamount to the false assumption that the purposes served by the Federal budget have a lower priority than those served in other ways. That kind of assumption is the height of folly, once we properly define the Federal budget as the greatest single identification of—and the greatest single instrument for—what we need to do together as a nation. If the pressure on our productive capabilities makes it desirable that we cut back somewhere, appropriate measures should be adopted to reduce or restrain other types of spending of a lower priority than the Federal budget, so that the great purposes of our national economic life may not be sacrificed while secondary purposes prevail.

A corollary of this position is that Federal spending should be projected on a long-range basis, covering a period of far more than one year. It should be advanced from year to year in absolute terms because of the growing needs of a growing population ; and because, if our economy grows adequately in the overall, it should not grow least in the things we need more of most. Even so, with an optimum rate of general economic growth, such Federal spending may decline as a percentage of total national production.

I believe that such a Federal spending policy would be a much wiser economic policy for growth, and also a much better counter-cyclical policy in the long run, than the application of the viewpoint that Federal spending should be adjusted upward and downward, depending upon shifts in general economic conditions. Such an upward and downward adjustment is irrational in terms of our national priorities, and is ineffective even as a stabilization device because it usually occurs too late or does not occur at all. It is true that, in the advent of a large depression, we might want to lift Federal spending even above the levels arrived at by the methods suggested above; but it is unlikely that we shall have to deal with this problem in the foreseeable future.

Taxation. Federal taxation represents an entirely different problem. Taxation, unlike public spending for needed purposes, has little or no positive value *per se*. While it is often said that taxation may help to improve income distribution, it is mainly the spending that puts income where it is needed ; the taxation is for other purposes. Thus, the real purpose of Federal taxation, after the level of Federal spending is appropriately determined, is to use taxation as a means of helping to achieve the desirable flow of income and purchasing power in facilitation of the three great purposes of our national economic life. If we are threatened with excessive pressures upon even optimum use of our resources (that is, inflationary pressures jeopardizing the maintenance of optimum growth and economic justice), then taxation should be high enough to run a budgetary surplus. It we suffer from large unemployment of plant and manpower, then taxation should be at levels which result in a deficit. In other words, it is tax policy, and not spending policy, which mainly should be used to deal with the stabilization problem.

I believe also that we should pay more attention to the composition or incidence of the tax burden, as distinguished from the total size of the tax take. Even from the viewpoint of combating those kinds of inflationary trends which can be dealt with by fiscal policy, I think that the nature of taxation is very important. A deficit arising from one composition of taxation may be less inflationary than a surplus arising from another composition of taxation. The same comment applies when the purpose of taxation is to stimulate rather than to restrain the economy. And the incidence of the tax burden has great bearing upon relative rates of investment and consumption, and also upon equity and economic justice. The aggregative approach to fiscal policy, with insufficient emphasis upon composition, is one of the gravest defects in our current economic thought and action.

Debt and surplus. One more point on fiscal policy. I believe that, if both Federal spending and Federal taxation were adjusted properly to the attainment of the three great purposes of our national economic life, the results in the long run would be Federal surpluses, providing some funds for reduction of the national debt. At the very least, we would average smaller Federal deficits than we have been averaging through the neglect of the three great purposes of our national economic life. This should be comforting to those who want to balance the budget and reduce the national debt, even though for my money I would be willing to run either surpluses or deficits if they were contributive to the optimum achievement of the three great purposes of our national economic life.

It is not entirely clear that the national debt, if well managed, is not a national asset, or that it is undesirable for the national debt to grow absolutely in a growing economy, even while it shrinks in proportion to the size of the total economy. Why, in a growing economy, is it desirable that only private debt should grow ? Certainly, in any event, we should gear fiscal policy to the three great purposes of our economic life, instead of conducting it in a vacuum, or in a space filled only with prejudices and preconceptions.

(4) *The same strictures apply even more vigorously to monetary policy. Our monetary policy in recent years has indeed been a farrago by any test.* Monetary policy, in recent years, has been used contrary to all of the three great purposes of our economic life ; and it has been inflationary to boot. The tighter monetary policy of 1952-53 was not needed to combat inflation, because price stability started in early or late 1951. But this tighter monetary policy did contribute to the recession which developed in 1954. Contrary to the frequent assertions that the economic upturn of 1954-55 was too great, and that monetary policy was too loose to curb it, this upturn did not carry us back to an optimum level of production ; and an excessively restrictive monetary policy contributed to a severely contracting rate of economic growth during the period 1955-57.*

---

*On into the 1960's, monetary policy has remained much too restrictive.

In early 1957, the tight money policy was defended on the ground that it was needed to restrain a classic excess of demand over available supplies, when in fact there was already far too much economic slack, and when some of us were pointing out that a recession was clearly on the way. Some of us also pointed out that the only kind of price inflation then in being was administered price inflation, and that the monetary policy was not suited to dealing with this kind of inflation. But the Federal Reserve Board insisted that there was really no distinction between administered price inflation and excessive demand inflation. The Federal Reserve Board also insisted, in early 1957, that the tight money policy was needed to induce consumers to save more so that more funds would be available for investment in plant and equipment, when in fact at that time the investment boom in plant and equipment was seriously outrunning private consumption and public demand. Meanwhile, total investment in early 1957, including investment in housing, State and local public investment in essential facilities and services, and investment opportunity on the part of small business, was running much too low to be consistent with the maintenance of the economic advance and the avoidance of recession. Thus, the tight money policy in early 1957, based upon a fundamental misreading of the economic situation, repressed the kind of investment which was already too low, repressed consumption which was already too low, and did nothing to restrain the boom in plant and equipment which was curbed only by the growing deficiencies in private and public consumption relative to productive capabilities. And the administered price increases were stopped not by the tight money policy; they were stopped belatedly by the most serious economic recession since World War II.

In early 1959, the Federal Reservo Board at long last started to talk about the evils of administered price inflation almost a year after this kind of inflation ceased, and finally admitted that the monetary policy was not suited to deal with this type of inflation. The Federal Reserve Board was actually saying that the monetary policy, if applied to this type of inflation, would only aggravate the distortions it creates. Nonetheless, the Federal Reserve System was again tightening up on money further, when there was still a tremendous slack in the economy; when unemployment was actually rising; when the States and localities were in dreadful difficulties with respect to financing their vital activities; when overall investment, including investment in plant and equipment was not expanding sufficiently to restore maximum prosperity; and when the only type of foreseeable inflation was a revival of administered price inflation in steel and other key industries long before they returned to reasonably full utilization of their productive capabilities.

To be sure, there is need for the selective and discriminating application of monetary policy. But the blunderbuss methods of the Federal Reserve System only aggravate the distortions in the credit and investment and income structure which are even more serious than the absolute levels, and in the overall repress production and employment although the urgent need is for expansion. Monetary policy so clearly divorced from the three great purposes of our national economic life needs thorough reconstruction, and instead of being independent needs to be vigorously integrated with the economic policies of the Executive branch. It also needs to be more thoroughly supervised by the Congress.

(5) *The final departure from basic economics is insufficient stress upon long-range policies and programs.* We are constantly seeking to make random and improvised adjustments to immediate situations. The idea seems to be that, if we take care of today, tomorrow will take care of itself. I think it is more nearly a fundamental truth that, if we take care of tomorrow, today will take care of itself. A family plans years ahead, a business plans decades ahead, and yet as a great nation we are trying to live from day to day. We cannot meet our domestic problems this way, and we certainly cannot meet the totalitarian challenge this way.

ELEMENTS OF A THEORY OF SYSTEM BEHAVIOUR

By OSKAR LANGE

Warsaw, Poland

1. INTRODUCTION

This essay is part of a larger work which is in preparation. The work is intended to provide a general theory of system behaviour. The theory is formulated in terms of cybernetic concepts. The present essay contains some of the basic concepts of this theory.

The wide range of Professor Mahalanobis's interests includes also philosophy. The problem of "Wholes and Parts" which is the subject of this essay, also, has an ancient tradition in Indian philosophic thought. It seems fit, therefore, to submit to this volume an essay which treats an ancient philosophical problem with the aid of mathematics. It shows the power of mathematical thought which the Editor of *Sankhyā* has successfully applied in a great variety of fields.

2. THE ACTIVE ELEMENT

By an active element E we mean a material object which depends in a given manner on other material objects and acts in a given manner on other material objects. The set of other material objects shall be called the environment of the given element.

We assume that:

(1) The environment acts on element E by inducing in it certain states of a strictly defined type such as temperature, pressure, electric charge, feeling, sense impression. The individual types of such states shall be called the inputs of element E.

(2) Element E acts on the environment by assuming certain states of a strictly defined character, e.g. temperature, magnetic field, colour, generation of sounds, motion. The individual types of such states shall be called the outputs of element E.

(3) Element E possesses at least one input and at least one output.

(4) The input states determine the output states uniquely.

Assumptions (1) and (2) are also expressed by the statement that element E is "relatively isolated."[1] This means that the element's contact with the environment takes place solely through the intermediary of its inputs and outputs. There are no other "channels" of contact. Assumption (3) excludes elements without inputs and outputs, i.e., monads. It also excludes elements with only inputs or only outputs, in other words, solely passive and solely active elements, so to speak. Each element through its inputs "receives" the action of the environment and through its outputs "transmits" action into the environment.

---

[1] This term was first introduced by Henryk Greniewski in *Elementy Logiki Indukcji* (*Elements of the Logic of Induction*), PWN, Warsaw, 1955, p. 28. Cf. also, Greniewski, *Cybernetics without Mathematics*, Pergamon Press, Oxford, PWN, Warsaw, 1960, pp. 9–10. Greniewski speaks of a "relatively isolated system." In view of the subject of the present paper, however, we prefer the term "active element," leaving the term "system" to denote a set of coupled elements.

Assumption (4) states the unique relation between the input and output states. We shall call this relation the *mode of action* of element E.

For the time being let us assume that the inputs and outputs of element E are finite in number. Let us denote the number of inputs by m and the number of outputs by n. Both m and n are natural numbers. In view of the Assumption (3) $m \geqslant 1$ and $n \geqslant 1$.
The difference

$$\delta = n - m \qquad \ldots \quad (2.1)$$

will be called the *diversification* of element E. Obviously δ is an integer. It may happen, however, that $\delta > 0$ as well as $\delta < 0$ or $\delta = 0$.

The input and output states of element E can be represented by numbers. If the input or output state consists of certain attributes (e.g., red colour, a given sound or movement) appearing or not, then we denote the state by 1 when the attribute appears and 0 when it does not. If the input or output state has the character of a discretely varying magnitude (e.g., number of particles per second, number of flies in an experimental jar), we denote the state by rational number that serves as a measure of that quantity. If, finally, the input or output state is a magnitude varying continuously (e.g., temperature, weight, height) we denote this state by a real number.

In this way, numbers—let us denote them by $x_1, x_2, \ldots, x_m$—can be assigned to the inputs to express their individual states. These numbers form a vector

$$x = (x_1, x_2, \ldots, x_m), \qquad \ldots \quad (2.2)$$

which represents the state of inputs. We call it the *input vector* and the numbers $x_1, x_2, \ldots, x_m$ constitute its components.

In similar fashion, the states of individual outputs are assigned numbers which we denote by $y_1, y_2, \ldots, y_n$. These are the components of the *output vector*.

$$y = (y_1, y_2, \ldots, y_n). \qquad \ldots \quad (2.3)$$

Both vectors, the input and output, are subject to the rules of vector algebra.

The active element E is represented graphically as follows :

Fig. 1.

In this drawing, on the left the input vector is indicated by an arrow and the output vector is on the right, also in the form of an arrow. Sometimes it is useful to show all the components of both vectors on the drawing. To avoid confusion we denote the components with broken lines. Then the drawing assumes the form below:

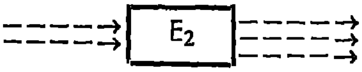


Fig. 2.

In Figure 2 the input vector has two components and the output vector has three.

The relation between the input and output states, in other words, the mode of action of the element E, can be expressed mathematically as a *transformation* of the vector x into the vector y. Symbolically, we write this transformation

$$y = T(x). \qquad \dots (2.4)$$

The set of admissible values of the vector x is called the domain of the transformation; the set of admissible values of the vector y is called the field of the transformation. The symbol T is called the transformation operator; it expresses the rule according to which vector x is transformed into vector y.

This rule can be expressed by means of a certain matrix. For this purpose, we denote by Δx_j the change in the value of the j-th component of vector x and by Δy_i the change in the value of the i-th component of vector y. Let us assume that the j-th component of the vector x changes by Δx_j and the other components of this vector remain unchanged. Then, in accordance with the transformation (2.4) there is a specified change Δy_i in the i-th component of the vector y.

The ratio

$$a_{ij} = \left(\frac{\Delta y_i}{\Delta x_j} \right)_{\Delta x_k = 0 \text{ for } k \neq j} \qquad \left(\begin{matrix} i = 1, 2, \dots, n \\ j = 1, 2, \dots, m \end{matrix} \right) \qquad \dots (2.5)$$

is a measure of the partial effect of a change in the j-th component of the vector x on the i-th component of the vector y. Let us call it the *coefficient of partial effect*.

The coefficients of partial effect form a matrix of n rows and m columns which we shall denote by A. We then have

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}. \qquad \dots (2.6)$$

We call this matrix the *transformation matrix*.

With the aid of the transformation matrix the rule for the transformation of the vector x into the vector y can be expressed in the form of a system of equations, namely;

$$\left. \begin{aligned} \Delta y_1 &= a_{11}\, \Delta x_1 + a_{12}\, \Delta x_2 + \dots + a_{1m}\, \Delta x_m, \\ \Delta y_2 &= a_{21}\, \Delta x_1 + a_{22}\, \Delta x_2 + \dots + a_{2m}\, \Delta x_m, \\ &\dots \quad \dots \quad \dots \\ \Delta y_n &= a_{n1}\, \Delta x_1 + a_{n2}\, \Delta x_2 + \dots + a_{nm}\, \Delta x_m. \end{aligned} \right\} \qquad \dots (2.7)$$

This can be expressed even more simply in the form of a vector equation

$$\Delta y = A\Delta x, \qquad \ldots (2.8)$$

where Δx and Δy are vectors, viz., $\Delta x = (\Delta x_1, \Delta x_2, \ldots, \Delta x_m)$ and $\Delta y = (\Delta y_1, \Delta y_2, \ldots, \Delta y_n)$. As can be seen, the transformation rule, which is symbolised by the operator T, is represented here explicitly in the form of the algebraic operations constituting the transformation in question.

The expression of the transformation rule in the form of (2.7) or (2.8) shall be henceforth called the *difference form* of this rule. If the coefficients of partial effect a_{ij} are constant, we immediately obtain the *integral form* of the transformation rule. Namely,

$$\left.\begin{aligned} y_1 &= a_{11}x_1+a_{12}x_2+\ldots+a_{1m}x_m, \\ y_2 &= a_{21}x_1+a_{22}x_2+\ldots+a_{2m}x_m, \\ &\ldots\quad \ldots\quad \ldots\quad \ldots\quad \ldots \\ y_n &= a_{n1}x_1+a_{n2}x_2+\ldots+a_{nm}x_m \end{aligned}\right\} \qquad \ldots (2.9)$$

or, more simply,

$$y = Ax. \qquad \ldots (2.10)$$

On the right hand side of these two expressions, i.e. (2.9) and (2.10), arbitrary constants may be added. For the sake of convenience we take these to be equal to zero.

The difference form shows the relation between the changes in the vector x and changes in the vector y. The integral form shows the relation between the values of the vectors x and y. When the coefficients a_{ij} are constant and, therefore, the integral form of the transformation has the form of (2.9) or (2.10) we say that the transformation is *linear*.

When coefficients a_{ij} are not constant they are functions of the vector x. The transformation matrix A is then a functional matrix. If the vector x is a continuous variable, (2.7) or (2.8) are differential equations. The solution of these differential equations, if there is one, consists of the system of functions

$$\left.\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_m), \\ y_2 &= f_2(x_1, x_2, \ldots, x_m), \\ &\ldots\quad \ldots\quad \ldots \\ y_n &= f_n(x_1, x_2, \ldots, x_m) \end{aligned}\right\} \qquad \ldots (2.11)$$

such that

$$a_{ij} = \frac{\partial f_i}{\partial x_j} \quad (i = 1, 2, \ldots, n;\; j = 1, 2, \ldots, m) \qquad \ldots (2.12)$$

This system can be written more simply in the form of a vector function :

$$y = f(x). \qquad \ldots (2.13)$$

The functions (2.11) or (2.13) constitute the integral form of the rule for the transformation of the vector x into the vector y.

The result (2.11) or (2.13) is also obtained in the case when the vector x varies in a discrete way. (2.7) or (2.8) are then difference equations and the functions (2.11) or (2.13) are the solutions of those equations.

The transformation (2.4) of the vector x into the vector y permits the state of the outputs of element E to be determined from the state of the inputs. Sometimes it is possible to make the inverse transformation

$$x = T^{-1}(y). \qquad \ldots \quad (2.14)$$

The inverse transformation permits conclusions as to the state of the inputs of element E to be drawn from the state of its outputs. An instance of such inverse transformation is a medical diagnosis of the causes of an ailment on the basis of its symptoms, or else the deciphering of a coded message.

The inverse transformation reduces to the solution of (2.9) or (2.11) or, which amounts to the same thing, of (2.10) or (2.13), with respect to the vector x, given the value of the vector y. As we know from algebra, the possibility of inverse transformation depends on two conditions. Firstly, we must have $n \geqslant m$, for otherwise there would be fewer equations than unknowns (the components of vector x are unknowns and there are m of them). In other words, the diversification of the element E cannot be negative since in that case there would not be enough outputs to draw conclusions as to the state of the inputs. Secondly, the rank of the transformation matrix must be m. In other words, exactly m partial effects of changes in the state of the individual inputs must be linearly independent.

If the inverse transformation is possible, the transformation rule can be written

$$x = A^{-1}y \qquad \ldots \quad (2.15)$$

for linear transformation, and

$$x = f^{-1}(y) \qquad \ldots \quad (2.16)$$

for the general case. Here, A^{-1} is the inverse of the matrix A and $f^{-1}(y)$ is the inverse vector function of the vector function $f(x)$.

3. COUPLING OF ACTIVE ELEMENTS

In accordance with Assumptions (1) and (2) of the preceding section, elements can interact only through their outputs and inputs. Let E_1 and E_2 be two active elements, $x^{(1)}$ and $x^{(2)}$ their input vectors and $y^{(1)}$ and $y^{(2)}$ their output vectors. Element E_1 can act on element E_2 only in such a way that E_2 "receive" as inputs the state of (all or some) outputs of E_1. For instance, element E_2 "receives" as inputs such output states of element E_1 as temperature and colour but not sounds. In other words, certain components of the output vector $y^{(1)}$ become components of the input vector $x^{(2)}$.

Denoting the components of the vector $y^{(1)}$ by $y_1^{(1)}, y_2^{(1)}, \ldots, y_n^{(1)}$ and the components of the vector $x^{(2)}$ by $x_1^{(2)}, x_2^{(2)}, \ldots, x_m^{(2)}$, we can write

$$x_j^{(2)} = y_i^{(1)} \qquad \ldots \quad (3.1)$$

for certain values of i and j, where $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$. Of course, (3.1) must be satisfied for at least one pair of values (i, j), for otherwise there would be no action by element E_1 on element E_2.

When components of the output vector of element E_1 are thus transformed into components of the input vector of element E_2, the two elements are said to be *coupled*.

The coupling of element E_1 with element E_2 is represented graphically as follows :

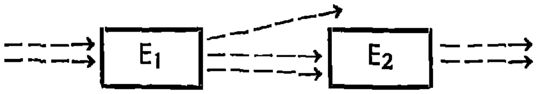


Fig. 3a.

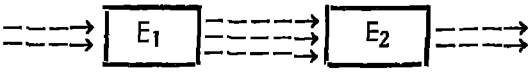


Fig. 3b.

In both the figures all of the components of the input and output vectors are shown. In Figure 3b element E_1 is coupled with element E_2 through all the components of its output vector whereas in Figure 3a, only through two components of that vector.

In practice it is rather inconvenient to use the individual components of the output and input vectors. With a large number of components the graphic picture of the coupling of elements becomes complicated. This picture, however, can be simplified in the following way. We introduce a square matrix which we denote by s_{12}, consisting of m or n rows and columns, depending on whether $m > n$ or $n > m$. The elements of the matrix have a value of 1 for all i and j for which (3.1) is satisfied and they have a value of 0 for all i and j for which the equality is not satisfied. In other words, an element of the matrix has a value of 1 if the corresponding component of vector $y^{(1)}$ becomes a component of vector $x^{(2)}$ and a value of 0 if this is not the case. The matrix introduced, has, therefore, the following shape (for example)

$$s_{12} = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 1 & 0 & \dots & 0 \end{bmatrix}. \qquad \dots (3.2)$$

This is a zero-one matrix with at least one element of the matrix being 1 and at most one element in each row (and column) equal to 1. If $m > n$, then $m-n$ rows consist of only zeros, and if $n < m$, then $n-m$ columns consist of only zeros; in that case $m-n$ inputs of element E_2 or $n-m$ outputs of element E_1 do not "participate" in the coupling. Matrix s_{12} shall be called the *coupling matrix* of element E_1 with element E_2.

With the aid of the coupling matrix, (3.1) can be rewritten in the form

$$x^{(2)} = s_{12}y^{(1)}, \qquad \dots (3.3)$$

i.e. in the form of a vector equality.

The graphical picture of the coupling of the element E_1 with the element E_2 can be simplified by drawing as follows :

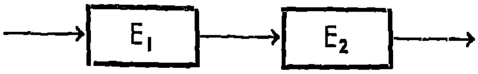


Fig. 4.

This figure shows only vectors and not their individual components. The arrow denoting the coupling of the element E_1 with the element E_2 stands for the vector equality (3.3).

Now let us consider the coupling of a larger number of active elements. We denote the individual elements by $E_1, E_2, E_3, \ldots$, and their input and output vectors, respectively, by $x^{(1)}, x^{(2)}, x^{(3)}, \ldots ; y^{(1)}, y^{(2)}, y^{(3)}, \ldots$.

If all the elements are coupled in pairs, then we say that there is a *chain of couplings*. Since the numeration of the elements is arbitrary, the chain of couplings can be represented as a sequence of vector equalities

$$\begin{aligned} x^{(2)} &= s_{12} y^{(1)}, \\ x^{(3)} &= s_{23} y^{(2)}, \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{aligned} \qquad \ldots \quad (3.4)$$

Here $s_{12}, s_{23}, \ldots$ denote the coupling matrices of element E_1 with element E_2, element E_2 with element E_3, etc. A vector equality in the sequence of (3.4) is called a *link* in the chain of couplings.

A chain of couplings is represented graphically as follows :

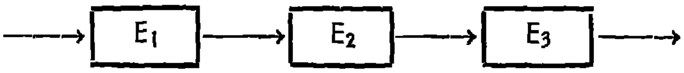


Fig. 5.

We differentiate between *open* and *closed* chains of couplings, called also open and closed *loops*. In a closed chain (loop) some active element is coupled with an element preceding it in the chain. Such coupling with an element appearing earlier in the chain of couplings is known as *feed-back*. A feature of a closed chain of couplings is that it has a feed-back among its links. A chain which does not have feed-back among its links is said to be open (an open loop).

Since the numeration of the elements is arbitrary, a closed chain of couplings can be expressed in the form of the vector equalities

$$\begin{aligned} x^{(2)} &= s_{12}\, y^{(1)}, \\ x^{(3)} &= s_{23}\, y^{(2)}, \\ \ldots \ &\ldots \ \ldots \\ x^{(1)} &= s_{r1}\, y^{(r)}. \end{aligned} \qquad \ldots \quad (3.5)$$

The last vector equality in this case is a feed-back. A feed-back occurs between element E_r and element E_1 (of course, we assume that $r > 1$).

A graphic picture of an open chain of couplings is given in Figure 5. A closed chain is depicted (by an example) in Figure 6. In this figure, there is a feed-back between element E_3 and element E_1.

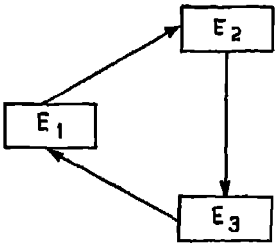


Fig. 6.

An active element can be coupled with more than one other active element. Conversely, more than one element can be coupled with a given element. We say then that the couplings are *ramified*, there being *ramification of output* in one case and *ramification of input* in the other.

In Figure 7a we give an example of ramification of output and in Figure 7b we give an example of ramification of input.

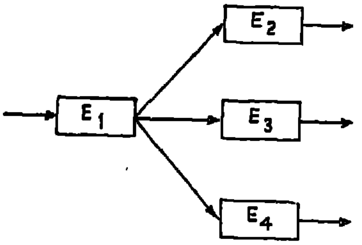


Fig. 7a.

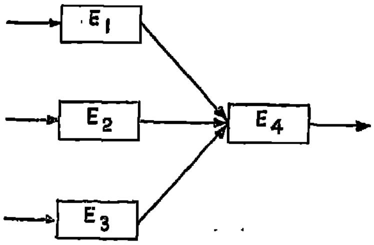


Fig. 7b.

Ramification of output is expressed by the following vector equalities.

$$\begin{aligned} x^{(2)} &= s_{12}\, y^{(1)}, \\ x^{(3)} &= s_{13}\, y^{(1)}, \\ \ldots &\quad \ldots \quad \ldots \\ x^{(q+1)} &= s_{1,q+1}\, y^{(1)}. \end{aligned} \qquad \ldots \quad (3.6)$$

Here, q denotes the number of elements coupled with element E_1.

In similar way the ramification of input is expressed by the vector equalities

$$\begin{aligned} x^{(q+1)} &= s_{1,q+1}\, y^{(1)}, \\ x^{(q+1)} &= s_{2,q+1}\, y^{(2)}, \\ \ldots &\ldots\ldots\ldots \\ x^{(q+1)} &= s_{q,q+1}\, y^{(q)}. \end{aligned} \qquad \ldots \quad (3.7)$$

In this case, q stands for the number of elements with which element E_{q+1} is coupled.

In a set containing a larger number of coupled active elements, there can be many different arrangements of couplings: open and closed chains (loops) and all sorts of ramifications. A set of such varied arrangements of couplings is called a *network* of couplings. A network of couplings can be expressed as a set of vector equalities of the type of (3.4), (3.5), (3.6) and (3.7). It can also be represented graphically.

A graphic picture of a network of couplings is given in Figure 8 :

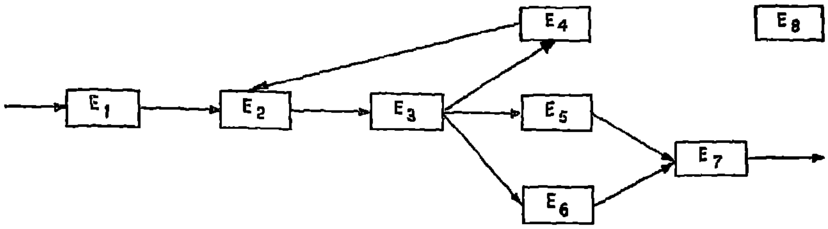


Fig. 8.

In this example there is an open chain (loop) of couplings between elements E_1, E_2 and E_3, a closed chain (loop) of couplings between elements E_2, E_3 and E_4, and a ramification of output at element E_3 and a ramification of input at element E_7. Finally, element E_8 is not coupled with any element in the given set of acting elements. All this arrangement of couplings (and absence of couplings) constitutes a network of couplings as shown in Figure 8.

Networks of couplings are a special case of so-called *graphs*, i.e. of abstract relations between the elements of a set. They are treated by a separate branch of mathematics known as the theory of graphs.[2] This theory is useful in investigation of the properties of networks of couplings between active elements.

4. THE SYSTEM OF ACTIVE ELEMENTS AND ITS STRUCTURE

A set of coupled active elements is called a system of active elements or, briefly a *system*.[3] Each element of a system is coupled with at least one other element of the system or at least one other element in the system is coupled with the given element. In a system, therefore, there are no isolated elements, i.e. elements which are neither coupled with any other element nor have any other element coupled with them.

The network of couplings between elements of a system is called the *structure* of the system.

We assume that a system contains a finite number of elements. The number of elements in the system is denoted by N. The individual elements are denoted by $E_1, E_2, \ldots, E_N$. The input and output vectors are denoted by $x^{(1)}, x^{(2)}, \ldots, x^{(N)}$ and $y^{(1)}, y^{(2)}, \ldots, y^{(N)}$, respectively.

---

[2] The systematic mathematical study of graphs was initiated by Deneš Koenig, *Theorie der endlichen und unendlichen Graphen*, Leipzig, 1936. Koenig introduced the term "graph." An exposition of the subject as it stands today is given by Claude Berge, *Théorie des graphes et ses applications*, Paris, 1958.

[3] Generally speaking, "system" is taken to mean a set of elements together with the set of relations between the elements. The set of such relations (and of all their isomorphic transformations) is called the *structure* of the system. Cf. for instance, Paul Lorenzen, *Einfuehrung in die operative Logik und Mathematik*, Berlin 1955, pp. 240–241. In our case, we have a set of active elements, the couplings of these elements are the relations.

If an element E_r is coupled with element E_s, the vector equality

$$x^{(s)} = s_{rs}\, y^{(r)} \qquad \ldots \quad (4.1)$$

is satisfied. In this equality, s_{rs} is the coupling matrix of element E_r with element E_s. As we know,[4] this is a square zero-one matrix in which at least one element is 1 and every row and every column contains at most one element equal to 1.

We now generalize this matrix for all elements E_r and E_s regardless of whether E_r is coupled with E_s or not. This is done by writing, for the case when element E_r is not coupled with element E_s,

$$s_{rs} = 0. \qquad \ldots \quad (4.2)$$

In other words, in this case we define the coupling matrix as a zero matrix (i.e. one in which all the elements are zero). In this case when the active element E_r is not coupled with the active element E_s, therefore, the vector equality (4.1) takes on the form

$$0 = s_{rs}\, y^{(r)}. \qquad \ldots \quad (4.3)$$

Having generalized the coupling matrix s_{rs} in this way, we can write the vector equality (4.1) for every pair of elements E_r and E_s, regardless of whether E_r is coupled with E_s or not. We then have $N(N-1)$ vector equalities

$$x^{(s)} = s_{rs}\, y^{(r)} \quad (r, s = 1, 2, \ldots, N; \quad r \neq s). \qquad \ldots \quad (4.4)$$

We assume that $r \neq s$ since matrix s_{rs} is always a zero matrix for $r = s$. By definition the active element is coupled with other active elements. No element is coupled with itself.

The coupling matrices, thus generalized, together form a square matrix with N rows and N columns. Namely,

$$S = \begin{bmatrix} 0 & s_{12} & \cdots & s_{1N} \\ s_2 & 0 & \cdots & s_{2N} \\ .\,. & .\,. & \cdots & .\,. \\ s_{N1} & s_{N2} & \cdots & 0 \end{bmatrix}. \qquad \ldots \quad (4.5)$$

This matrix, which we denote by S, shall be called the *structure matrix* of the system. For it represents the structure of the system, i.e. the network of couplings of the elements of the system.

The matrix S is a zero-one matrix because all coupling matrices s_{rs}, which are its sub-matrices, are zero-one matrices.

Since each element in the system is coupled with at least one other element or at least one other element is coupled with the given element, the minimum number of couplings in the system is therefore, $N-1$. The matrix, therefore, contains at least

---

[4] Cf. (3.3).

[5] Cf. (3.2).

$N-1$ non-zero submatrices s_{rs}. Moreover, these non-zero sub-matrices are so arranged that if the r-th row of the matrix S contains only zero sub-matrices, the r-th column must contain at least one non-zero sub-matrix. For if the element E_r is not coupled with any other element then at least one other element must be coupled with the element E_r.

An element which is not coupled with any other element in the system or which has no other element coupled to it is called a *boundary element* of the system. An element in the system which is not a boundary element shall be called an *interior element* of the system. The set of boundary elements of a system shall be called the system's *surface* and the set of the interior elements of the system shall be called as the system's *interior*. For the sake of illustration let us look at Figure 8 above.

The active elements E_1, E_2, E_3, E_4, E_5, E_6 and E_7 form a system; the isolated element E_8 does not belong to it. Elements E_1 and E_7 are boundary elements of the system; the set of them constitutes the surface of the system. On the other hand, elements E_2, E_3, E_4, E_5 and E_6 are interior elements of the system; the set of them constitutes the interior of the system.

Among the boundary elements of the system we distinguish between those which are not coupled with other elements in the system and those with which no other elements in the system are coupled. In the case of the first-mentioned, the components of the output vector are not components of the input vector of any other element of the system. We shall call them the output boundary elements. In the case of the last-mentioned, the components of the input vector are not components of the output vector of any other element in the system. We shall call them the input boundary elements. In Figure 8, E_1 is an input boundary element whereas E_7 is an output boundary element.

The input and output boundary elements are expressed in the structure matrix of the system. If element E_r is an input boundary element, then the r-th column of the matrix S contains only zero sub-matrices. For, in keeping with the generalized definition of the couplings matrix, we then have

$$s_{1r} = s_{2r} = \dots = s_{Nr} = 0.$$

If, on the other hand, element E_r is an output boundary element, then in accordance with the definition we have

$$s_{r1} = s_{r2} = \dots = s_{rN} = 0.$$

In this case the r-th row of the matrix S consists of only zero sub-matrices.

Of course, the converse is also true; a column in the matrix S consisting of only zero sub-matrices signifies that the corresponding element is an input boundary element and a row of the matrix S consisting of only zero submatrices signifies that the corresponding element is an output boundary element. It follows that the set of all the columns and rows of the matrix S consisting of only zero sub-matrices corresponds to the system's surface. On the other hand, the columns and rows containing at least one non-zero sub-matrix s_{rs} correspond to the interior of the system.

A system which does not contain any boundary elements, that is, does not possess a surface, shall be called a *closed system.* From what has been said it follows that the structure matrix of a closed system contains neither columns nor rows consisting of only zero sub-matrices; each column and each row contains at least one non-zero sub-matrix. In a closed system, therefore, there are at least N couplings. A system which contains boundary elements shall be called an *open system.* The structure matrix of an open system contains columns or rows (or both) consisting of only zero sub-matrices and representing together the surface of the system.

The structure matrix of the system reflects also other properties of the couplings of the elements of the system. Feed-backs, for instance, are all represented by sub-matrices s_{rs} which lie below the diagonal of the matrix. Feed-backs are couplings with an element preceding the given element in the chain of couplings. The corresponding coupling matrices s_{rs} have $s < r$. This determines their position below the diagonal.

A closed system possesses at least one feed-back. Let us assume that there are no feed-backs in the system. Then the matrix of the system is

$$S = \begin{pmatrix} 0 & s_{12} & \dots & s_{1N} \\ 0 & 0 & \dots & s_{2N} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}. \qquad \dots \quad (4.6)$$

As can be seen, there is then at least one input boundary element and at least one output boundary element. No element is coupled with element E_1 and element E_N is not coupled with any element (we recall that the numeration of the elements is arbitrary). The system, therefore, cannot be a closed one.

An open chain of couplings (open loop) is expressed by a structure matrix of the system in which all sub-matrices s_{rs} below the diagonal are zero and, moreover, in each row (apart from the last one which consists entirely of zero sub-matrices) as well as in each column (apart from the first one which consists entirely of zero sub-matrices) there is exactly one non-zero sub-matrix s_{rs}. In other words, the structure matrix of the system has the following appearance :

$$S = \begin{bmatrix} 0 & s_{12} & 0 & \dots & 0 & 0 \\ 0 & 0 & s_{23} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & s_{N-1,N} \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \qquad \dots \quad (4.7)$$

If the chain of couplings is closed (closed loop), then in (4.7) there is at least one non-zero matrix below the diagonal.

If a row of the structure matrix of the system contains more than one non-zero sub-matrix s_{rs}, this indicates a ramification of output. For example, let there be non-zero sub-matrices s_{12}, s_{13}, s_{14}, in the first row of the matrix S; then element E_1 is coupled with elements E_2, E_3 and E_4. Similarly, if a column of the structure matrix contains more than one non-zero sub-matrix s_{rs}, this is an indication of a ramification of input. For example, in matrix (4.6) if s_{1N} and s_{2N} are non-zero, element E_1 and element E_2 are coupled with element E_N.

Thus properties of the network of couplings of the system can be read off from the system's structure matrix.

5. THE MODE OF ACTION OF A SYSTEM

The inputs and outputs of a system containing N active elements, satisfy $N(N-1)$ vector equalities (4.4), i.e.

$$x^{(s)} = s_{rs}\, y^{(r)} \quad (r, s = 1, 2, \ldots, N;\ \ r \neq s). \qquad \ldots \quad (5.1)$$

These equalities represent the couplings of the elements of the system.

Each element of the system possesses a definite mode of action expressed mathematically by the transformation (2.4). Denoting by T_r the transformation expressing the mode of action of element E_r, we have N such transformations :

$$y^{(r)} = T_r(x^{(r)}) \qquad (r = 1, 2, \ldots, N). \qquad \ldots \quad (5.2)$$

Substituting the transformation (5.2) into the vector equalities (5.1), we obtain

$$x^{(s)} = s_{rs} T_r(x^{(r)}) \qquad (r, s = 1, 2, \ldots, N;\ \ r \neq s). \qquad \ldots \quad (5.3)$$

This is a set of $N(N-1)$ transformations.

Writing

$$R_{rs} = s_{rs} T_r, \qquad \ldots \quad (5.4)$$

we can present these transformations symbolically as

$$x^{(s)} = R_{rs}(x^{(r)}) \qquad (r, s = 1, 2, \ldots, N;\ \ r \neq s). \qquad \ldots \quad (5.5)$$

Subjecting the vector equalities (5.1) to transformations (5.2), we obtain

$$T_s(x^{(s)}) = T_s s_{rs}(y^{(r)}),$$

and consequently

$$y^{(s)} = T_s s_{rs}(y^{(r)}) \quad (r, s = 1, 2, \ldots, N;\ \ r \neq s). \qquad \ldots \quad (5.6)$$

We again obtain a set of $N(N-1)$ transformations. Writing

$$P_{rs} = T_s s_{rs}, \qquad \ldots \quad (5.7)$$

we can express these transformations symbolically in the form

$$y^{(s)} = P_{rs}(y^{(r)}) \qquad (r, s = 1, 2, \ldots, N;\ \ r \neq s). \qquad \ldots \quad (5.8)$$

The set of transformations (5.5) assigns new values of the input vectors of the system's elements to the given initial values of these vectors. Similarly, the set of transformations (5.8) assigns new values of the output vectors of the system's elements to the given initial values of these vectors. The sets of transformations (5.5) and (5.8) express, as we say, the *mode of action of the system.* They indicate in what manner a set of given states of the inputs and outputs of the elements of a system is transformed into a new set of states of those inputs and outputs. To employ an expression of Karl Marx, it can be said that the two sets of transformations express "*the inner law of motion of the system.*"[6]

---

[6] Cf. Karl Marx, *Capital*, Vol. I, Foreign Languages Publishing House, Moscow, 1954, p. 10.

We now introduce the *vector of the input states of the system* as a compound vector, its components being the input vectors of the elements belonging to the system. Denoting this vector by X, we have

$$X = (x^{(1)}, x^{(2)}, \dots, x^{(N)}). \qquad \dots \quad (5.9)$$

We introduce in a similar way a compound *vector of the output states of the system.* We denote this vector by Y and write

$$Y = (y^{(1)}, y^{(2)}, \dots, y^{(N)}). \qquad \dots \quad (5.10)$$

These compound vectors are equivalent to quasi-diagonal matrices[7]

$$X = \begin{pmatrix} x^{(1)} & 0 & \dots & 0 \\ 0 & x^{(2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x^{(N)} \end{pmatrix} \qquad \dots \quad (5.9a)$$

and

$$Y = \begin{pmatrix} y^{(1)} & 0 & \dots & 0 \\ 0 & y^{(2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & y^{(N)} \end{pmatrix}. \qquad \dots \quad (5.10a)$$

Denoting by m_r the number of inputs and by n_r the number of outputs of element E_r $(r = 1, 2, \dots, N)$, we find that the matrix X has N rows and $\sum_{r=1}^{N} m_r$ columns and the matrix Y has N rows and $\sum_{r=1}^{N} n_r$ columns. As need be, X and Y can be interpreted as either compound vectors (5.9) and (5.10) or as quasi-diagonal matrices (5.9a) and (5.10a).

The set of transformations (5.5) is equivalent to the following matrix equality

$$\begin{bmatrix} 0 & x^{(2)} & \dots & x^{(N)} \\ x^{(1)} & 0 & \dots & x^{(N)} \\ \dots & \dots & \dots & \dots \\ x^{(1)} & x^{(2)} & \dots & 0 \end{bmatrix} = \begin{bmatrix} 0 & R_{12}(x^{(1)}) & \dots & R_{1N}(x^{(1)}) \\ R_{21}(x^{(2)}) & 0 & \dots & R_{2N}(x^{(2)}) \\ \dots & \dots & \dots & \dots \\ R_{N1}(x^{(N)}) & R_{N2}(x^{(N)}) & \dots & 0 \end{bmatrix}. \qquad \dots \quad (5.11)$$

Similarly, the set of transformations (5.8) is equivalent to the matrix equality

$$\begin{bmatrix} 0 & y^{(2)} & \dots & y^{(N)} \\ y^{(1)} & 0 & \dots & y^{(N)} \\ \dots & \dots & \dots & \dots \\ y^{(1)} & y^{(2)} & \dots & 0 \end{bmatrix} = \begin{bmatrix} 0 & P_{12}(y^{(1)}) & \dots & P_{1N}(y^{(1)}) \\ P_{21}(y^{(2)}) & 0 & \dots & P_{2N}(y^{(2)}) \\ \dots & \dots & \dots & \dots \\ P_{N1}(y^{(N)}) & P_{N2}(y^{(N)}) & \dots & 0 \end{bmatrix}. \qquad \dots \quad (5.12)$$

---

[7] A quasi-diagonal matrix is one in which sub-matrices or vectors appear along the diagonal and all the remaining elements are equal to zero. For a definition of a quasi-diagonal matrix see V. I. Smirnov, *Kurs Vyshei Matematiki*, Vol. III, Part 1, 4th ed., Leningrad—Moscow, 1949, p. 104.

From these equalities it follows that the initial state of the input vectors $x^{(1)}, x^{(2)}, \ldots, x^{(N)}$ uniquely determines the state of these vectors after transformation. In the same way, the initial state of the output vectors $y^{(1)}, y^{(2)}, \ldots, y^{(N)}$ uniquely determines the state of these vectors after transformation. With the aid of the compound vectors (5.9) and (5.10) we can express this as a transformation of these vectors. Instead of the matrix equality (5.11) we write

$$X' = R(X), \quad \ldots \quad (5.13)$$

and in place of the matrix equality (5.12),

$$Y' = P(Y). \quad \ldots \quad (5.14)$$

Here, R and P are transformation operations, X and Y represent the initial state of the inputs and outputs of all the elements of the system, whereas X' and Y' denote the new state of the inputs and outputs after the transformation.

It suffices in this case to consider either transformation (5.13) or transformation (5.14). For from the vector equalities (5.1) we obtain the matrix equality

$$\begin{bmatrix} 0 & x^{(2)} & \ldots & x^{(N)} \\ x^{(1)} & 0 & \ldots & x^{(N)} \\ \ldots & \ldots & \ldots & \ldots \\ x^{(1)} & x^{(2)} & \ldots & 0 \end{bmatrix} = \begin{bmatrix} 0 & s_{12}(y^{(1)}) & \ldots & s_{1N}(y^{(1)}) \\ s_{21}(y^{(2)}) & 0 & \ldots & s_{2N}(y^{(2)}) \\ \ldots & \ldots & \ldots & \ldots \\ s_{N1}(y^{(N)}) & s_{N2}(y^{(N)}) & \ldots & 0 \end{bmatrix}. \quad \ldots \quad (5.15)$$

The right-hand side of this equality can be written in the form YS where Y is the quasi-diagonal matrix (5.10a), equivalent to the compound vector (5.10), and S is the structure matrix of the system :

$$S = \begin{pmatrix} 0 & s_{12} & \ldots & s_{1N} \\ s_{21} & 0 & \ldots & s_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ s_{N1} & s_{N2} & \ldots & 0 \end{pmatrix}. \quad \ldots \quad (5.16)$$

The left-hand side of the above matrix equality is equivalent to the compound vector (5.9). As a result, we have

$$X = YS. \quad \ldots \quad (5.17)$$

The transformations (5.13) and (5.14), therefore, follow one from the other.

The operators R and P can be presented symbolically in the form of a matrix of operators appearing in the sets of transformations (5.5) and (5.8). Namely,

$$R = \begin{pmatrix} 0 & R_{12} & \ldots & R_{1N} \\ R_{21} & 0 & \ldots & R_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ R_{N1} & R_{N2} & \ldots & 0 \end{pmatrix} \quad \ldots \quad (5.18)$$

and

$$P = \begin{pmatrix} 0 & P_{12} & \ldots & P_{1N} \\ P_{21} & 0 & \ldots & P_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ P_{N1} & P_{N2} & \ldots & 0 \end{pmatrix}. \quad \ldots \quad (5.19)$$

Let us denote by T the diagonal matrix

$$T = \begin{pmatrix} T_1 & 0 & \dots & 0 \\ 0 & T_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & T_N \end{pmatrix}. \qquad \dots (5.20)$$

The non-zero elements of this matrix constitute the transformation operators corresponding to the individual elements of the system. This diagonal matrix shall be called the *matrix of modes of action of the elements.*

If (5.4) and (5.7) are taken into account, the operators R and P can be presented as products, viz.

$$R = TS \qquad \dots (5.21)$$

and

$$P = ST \qquad \dots (5.22)$$

where S is the structures matrix of the system (5.16) (cf. also (4.7) above).

In consequence, it shows that the system's mode of action, expressed mathematically by the transformation (5.11) or (5.12), can be presented in the form

$$X' = TS(X), \qquad \dots (5.23)$$

$$Y' = ST(Y). \qquad \dots (5.24)$$

As can be seen, the mode of action of the system depends both on T, the matrix of modes of action of the elements, as well as on the structure matrix S of the system. In order to determine the mode of action of the system it is not enough to know the modes of action of the elements, which are expressed mathematically by the matrix T, but it is necessary also to know the structure matrix S of the system, which represents the network of couplings of the elements.

A system of active elements, therefore, is endowed with its own specific mode of action, "a law of motion", consisting in the transformation of the input and output states of the elements. This mode of action is a feature of the system as a *whole,* for it depends not only on the mode of action of the individual elements but also on the structure of the system, i.e. on the network of couplings of the elements. This is expressed mathematically in (5.23) and (5.24). It is the structure represented by the matrix S that gives the system the quality of a whole. The same elements with the same modes of action but coupled in a different manner constitute a different system with a different mode of action of the system. In mathematical language : with an unaltered matrix T of modes of action of the elements, the mode of action of the system becomes different if there is a different structure matrix S of the system. A difference in the structure implies a difference in the system's mode of action.

In this way an explanation is provided for the question of wholes with laws of action which cannot be derived from the mere laws of action of the individual elements. Such wholes are simply sets of coupled active elements, i.e., systems. Their mode of action constitutes the combined result of the modes of action of the individual elements, and of the structure of the system, i.e. of the network of couplings of the system elements, of which the system is composed.

SCIENTIFIC APPROACH IN PLANNING

By M. MUKHERJEE

Indian Statistical Institute

1. INTRODUCTION

1.1. Thinking on national planning at a somewhat concrete level started quite early in India, and even in nineteen-twenties, that is, immediately after the Russian revolution, some intellectuals in India came upon the idea of national planning. Notably, M. Visvesvaraya saw quite early that conscious national economic planning was needed for rapid industrial development in India. Subsequently, he proceeded to elaborate a ten year plan, which was not unrealistic, except perhaps on the question of finance of the plan. The leadership of Jawaharlal Nehru as the Chairman of the National Planning Committee (NPC) set up in 1938 gave a great impetus to thinking on planning in the country, and some of the panel reports of the NPC bear close study even today. Though the NPC did not succeed in producing a blue-print of a national plan, it served two very important purposes: the intellectuals in the country got intimately acquainted with notions of planning and the most important political party in the country, to some extent, got committed to a programme of planning. Political upheavals and repressions cut short the work of the NPC. Public thinking and discussion on planning reached a new tempo in the post-war period because of the release of three important private investigations on the topic: the Gandhian Plan, the Peoples' Plan and the Bombay Plan. In so far as facts and figures presented are concerned, the Bombay Plan prepared by eight leading industrialists of India, was the most realistic. For example, the anticipated rate of growth of national income was 7 per cent in this plan: the other two plans anticipated double this rate. However, the Gandhian Plan based on the ideas of Mahatma Gandhi and the Peoples' Plan prepared under the auspices of the Indian Federation of Labour, put forward numerous illuminating ideas relating to several aspects of national planning. There were unhappy events in the country immediately before and after the independence and it took some time before Free-India could settle down and launch on planning. Even then, the Planning Commission was set up as early as in 1949 and was preceded by the Industrial Policy Resolution of 1948 which indicated certain ideas on planning. This brief sketch shows that the public opinion in the country was educated during the last thirty years so as to be able to receive the idea of national planning with appreciation. The traditional *laissez faire* policy of the Government of India during the British rule did not contribute towards rapid economic development of the country, and it was more or less clear to all sections of the educated public that India's chronic problem of poverty could not be solved unless we have recourse to national economic planning.[1]

1.2. But while the idea of planning was generally accepted, the word 'planning' connoted different notions to different groups of people in the early post-independence period. Even today we are still not very clear on this point though the understanding is getting better day by day. To some extent, this is unavoidable. Economic planning is closely linked with the national policy which is expected to set down our goals in economic, social and political fields and to indicate proximate descriptions of the paths we should follow to attain our goals. The policies should distinguish between our immediate goals from our ultimate goals, and cover not only our short period desiderata but also our more distant perspective. Likewise, spatial dimensions of development should be comprehended along with the temporal dimensions. It will be clearly seen why it is difficult to have unequivocal policy statements in this general field. Firstly, there must be clearly defined agreed goals and then there must be agreement about the path to be followed and measures to be taken for attaining the goals. The system has to cover all our important aspirations in economic, social and political fields. While we have succeeded in setting our tentative targets in so far as our goals are concerned, we are still partly divided about our notions in respect of the efficient paths and measures. It will be appreciated, on the other hand, that clear-cut and accepted policy pronouncements on both goals and paths would have helped us materially in attaining whatever objectives we have collectively set.

1.3. Our difficulties stem partly from our political traditions as well as from the basis of our political power and partly from the vestiges of traditional economic thinking which guided capitalist societies in the past. The basis of our political power as well as our political traditions allow us to be radical but not unduly so. We have accepted the procedure of discussion and persuasion even when we are sure of inherent justice of particular lines of action, knowing fully well that we are eschewing quicker methods and adopting slower ones. In the field of economics, while probably we have steered clear of the *laissez faire* doctrine, we are still very much wedded to Keynesian economics, which does not apply to Indian situation because the type of economy for which the system was developed was quite different from the Indian economy today.

1.4. In this atmosphere of somewhat vague thinking, inadequate formulation of basic issues, and relatively amorphous political alignments, Professor

---

[1] Mahalanobis (1959a) has summarised Nehru's contributions towards planning as follows : "Under the leadership of Nehru, India has made big advances. He initiated the thinking on planning in India. Through his speeches and through planning committees and the Planning Commission, he has exercised a profound educative influence and made India conscious of the need for national planning. Through the Congress Party and the Government, he has made planning an instrument of national policy on the biggest scale outside the communist countries, and has persuaded India to accept socialism as her goal. He has brought to Indian planning a full appreciation of the scientific revolution which is transforming the world, a sensitive awareness of human values and cultural traditions, an inherent sense of democracy and an international outlook."

Mahalanobis' advent has been a welcome event. As a scientist, he succeeded quickly in separating essentials from the non-essentials, and developing an interrelated programme of planning which is logically satisfying and unequivocally stated. In this process, he had to sacrifice many details, some of which may be of importance in the context of development planning. It is not possible, however, to comprehend all the complexities of a real economy in a logical system. The process of simplification, which is detrimental to richness of the content, is a necessary step. He was also handicapped by lack of data on several important factors and some of the simplifications were necessary because of this. This was unavoidable in an underdeveloped economy attempting to plan for the first time. In spite of serious difficulties, he succeeded in his pioneering effort towards a concrete formulation of a plan for economic development of India.

1.5. This essay is concerned with a brief descriptive account of his contributions in planning and allied subjects. A summarisation involves a selection of some elements or aspects of the totality of contributions. Since, such selection is essentially subjective, it is quite possible that, in the present write up, I have overplayed some less important aspects of his contribution and underplayed some more important aspects. This contingency, however, is unavoidable.

2. STATISTICS AND PLANNING

2.1. Mahalanobis has an uncanny knack of absorbing a mass of statistical material quickly and locating mistakes in calculations and errors of judgement in the logical scheme of analysis. He has, no doubt, acquired this faculty by not only having had experience of careful scrutiny of millions of figures in diverse branches of statistics but also from an intimate grasp of the methods used for computation of the figures. Thus apart from a mastery over the science, he has an acute diagnostic ability enabling him to spot the defects, as if intuitively. Perhaps it is this faculty which has made him a foremost statistician not only in India but also in the world today.

2.2. It is this diagnostic ability which came to his aid when he was drawn into a small group of thinkers on planning selected by Jawaharlal Nehru for shaping India's future development. He was able, within a very short time, to lay his finger on many of the defects from which planning in our country was suffering. But apart from this, he has contributed materially to planning in his capacity as a statistician, by ensuring that the planners get the facts and figures needed for planning. He had the foresight in realising that adequate and accurate statistics formed the basis of planning. Since decisions should be based on valid inferences, he has thought deeply on the relation between statistical errors and decisions based on estimates subject to errors. In his view, the notion of sampling error is often too narrow from this angle, and large errors may be permissible for certain kinds of decisions. Mahalanobis thus contributed to planning in two distinct fields. Firstly, as the leading statistician of the country, he has helped in getting together the bricks out of which plans are made of. And, secondly an accomplished diagnostician, he has gone beyond the facts and figures to recognise and isolate some of the graver defects in our planning and suggest measures for their remedy.

2.3. To some extent, of course, these two roles are mutually interacting, and one has to be conscious about this in trying to understand Mahalanobis' effective contribution to planning. How is it that a man who has not studied principles of economics in a systematic way could make a material contribution towards formulating the plan for economic development of the country? This baffling question can be answered only when we remember that as a statistician Mahalanobis was aware of the economic facts of the country better than many economists even before he started thinking on planning. He was, of course, not in favour of the refinements and niceties of deductively arrived economic theories developed in western countries to explain a situation different from ours. He would rather accept the facts about India than adopt or use theories obtained in this manner.[2]

2.4. While Mahalanobis was acquainted with Jawaharlal Nehru for a very long time, he first had long and intimate discussions on planning only in 1940 when Nehru wanted him to prepare a statistical commentary on the NPC reports. Subsequently Nehru maintained an active interest in the development of the Indian Statistical Institute (ISI).

2.5. After the Independence, the Prime Minister depended mostly on Mahalanobis' initiative to develop the organisation of statistics in the country. Mahalanobis started working as the Honorary Statistical Adviser to the Cabinet from February 1949, when a Central Statistical Unit was started in Delhi at his initiative. This unit was converted into the Central Statistical Organisation (CSO) two years later and was entrusted with the task of co-ordination of official statistics and laying down statistical standards. As Honorary Statistical Adviser to the Cabinet, Mahalanobis guided the activities of the CSO in all technical matters from its very inception and even today, this institution, which has grown rapidly, works under his technical guidance. It may also be noted that even earlier, Mahalanobis was instrumental in developing the State Statistical Bureau of West Bengal. The rapid development of official statistics in the country during the post-independence period is very much connected with the history of the CSO and the State Statistical Bureaus, and it may be noted that Mahalanobis is intimately connected with the CSO and also helped to set up as a model, the West Bengal State Statistical Bureau. While the development of official statistics in general is not directly linked with needs of planning, a sound general purpose statistical system is necessary for planning and Mahalanobis materially contributed in this direction.

---

[2] Some remarks by Mahalanobis on this topic are given below :

"I may start by saying that sophisticated economic theories, which may be appropriate for advanced countries, had acted for a long time as a formidable thought barrier to economic progress in India," Mahalanobis (1961b). Also, "The economic theory of highly developed countries appears to be basically static in character and they are concerned, above all, with the most efficient distribution of the stock of capital and other resources and not with the problems of economic development through an increase in capital accumulation. A theory of the same brand did not help the economic development of India or any other underdeveloped country; it hindered such developments on the other hand," (Mahalanobis, 1959c).

2.6. Mahalanobis was appointed the Chairman of the National Income Committee in 1949, and this event may be regarded as a corner-stone in the present context. Assimilation of diverse official statistics need not necessarily give a central or national view point to the practitioner. But thinking about national income, in a way, is thinking about the performance of the nation as a whole. While working as the Chairman of the Committee, Mahalanobis became acutely aware of the national problems, national resources and allied matters. It should be pointed out, however, that localised socio-economic surveys conducted by him earlier, including family budget surveys of industrial and rural workers and in particular, a survey of after-effects of the great famine in Bengal in 1943 made him acutely aware of our poverty long before he started thinking on national income. Also, even before this, Mahalanobis undertook some fundamental work in the fields of regional and partial planning. For example, his quantitative work on rainfall and floods have now flowered into the D.V. Project and Hirakud Project on Mahanadi.

2.7. The National Sample Survey (NSS) started at his initiative in 1950 was his answer to the problem of gaps in statistics. The survey is continuing even now and here we have an organisation which can furnish sample estimates of any character quickly depending on our need. In this sampling organisation, the CSO, the State Statistical Bureaus and the ISI collaborate with one another, and the whole scheme has been conceived with great vision and foresight. The data are collected from independent interpenetrating sub-samples by independent parties of investigators and processing centres are also geographically separated from one another, each centre taking a sub-sample or a part of a sub-sample. Thus, the whole system is conceived in such a way that investigation and processing errors and biases are controlled at all stages, and any data coming out of the system has a built-in guarantee of objectivity. It is extremely difficult to implement a gigantic scheme like this to perfection at the initial stage, but the result achieved in the past decade is encouraging.

2.8. Mahalanobis has maintained his interest in the NSS throughout its existence. He has not, however, been able to maintain close contact with the routine operations of the survey. On the other hand, he has made substantial use of NSS data for his scientific studies on planning and other problems.[3] While making such studies he has frequently become conscious about some of the deficiencies of the survey and tried to apply corrective steps to the extent possible. The general idea of the NSS comes from him, but in the actual conduct of the survey, his contributions have been partial and spasmodic.

2.9. The preceding paragraphs are intended to convey that Mahalanobis is closely connected with the development of operational statistics in the country during the post-independence period. The rapid development in the field probably owes more to his initiative than to that of any other single person. Secondly, even in this capacity, his work drew him to problems of national development. That is, the types of statistics he has to deal with required his being aware of the needs of planning in the country. Also, his awareness of the planning problems which he got from his association with the Planning Commission enabled him to develop the statistical system in the country keeping the planning needs as the principal objectives.

---

[3] As this essay is concerned with Mahalanobis' contribution to planning, we have not indicated his methodological contributions in the field of use of statistical data. The most important of these in the recent past is his novel idea of fractile graphical analysis. This gives a theoretically satisfying method of drawing valid inference from survey data of diverse types. In particular, changes in size distribution over time or over regions of some data given in terms of value could be tackled effectively by this method.

3. USE OF METHODS OF OPERATIONAL RESEARCH TO SOLVE PROBLEMS OF THE INDIAN SECOND FIVE YEAR PLAN

3.1. Mahalanobis has made use of some simple mathematical models for studying some aspects of the problems connected with planning. The models enabled him to isolate important sectors of the economy and study their mutual interdependence.[4] This part of the work can best be described as operational research in the field of planning, as he has himself done. The models developed do not follow the tenets of accepted economic theory and hence could not be included in traditional mathematical economics. On the other hand, as the models were developed in an atmosphere of pressing policy needs, much attention could not be paid to develop any refined methods for statistical estimation of the parameters involved. Also, available data in the country were inadequate for application of rigorous procedures of estimation. Consequently, this work should not probably fall under the scope of econometrics.

3.2. While developing the models, Mahalanobis' long experience as a professor of physics came handy. We have stressed so far his background in statistics and his attainments as an applied scientist and diagnostician and ventured to suggest that the concomitance of the two is a major reason for his success as a practical planner; his ability as a physicist, in addition, helped him to abstract and concretise his problems and ultimately arrive at reasonable solutions.

3.3. The simple concept of capital output ratio and the fact that the ability of a nation to step up its investment depends on the level of income was given by him in his presidential address to the 37th session of the Indian Science Congress (Mahalanobis, 1950). It will be recalled that in his capacity as Statistical Adviser to the Central Cabinet, he has been brought into contact with the activities of the Planning Commission early in 1949, and in less than one year, he has been able to identify and start thinking in terms of some of the basic variables. At the time, he was not in touch with the writings of economists on the topic, and obtained the notions independently. The concept of capital output ratio used by him was conceived as the ratio of investment to the aggregate output resulting directly from the investment activity. He also used an idea of development in all other activities as a consequence of investment in a particular activity and thus tacitly adopted ideas of inter-industrial economics.

---

[4] Mahalanobis' opinion about the use of models in this context is of considerable interest. "I do not think that the models have any permanent value of their own. I have used them as scaffolding to be dismantled as soon as their purpose has been served," (Mahalanobis, 1955b).

3.4. What is described in India as the single sector model of Mahalanobis is a forward-looking Harrod-Domar type of model. Mahalanobis presented these ideas first in a lecture delivered at the National Institute of Sciences of India in October 1952 (Mahalanobis, 1952). The variables involved in the model are national income (Y_t) and net investment (I_t)[5], where the subscript t stands for time. The two variables are connected by the relations

$$\frac{I_t}{Y_t} = \alpha \qquad \dots (1)$$

$$\frac{Y_{t+1} - Y_t}{I_t} = \beta \qquad \dots (2)$$

where α and β, parameters to be estimated, are described respectively by him as the rate of investment and the income coefficient of investment. The two equations taken together gives the growth path of national income as

$$Y_t = Y_0(1+\alpha\beta)^t \qquad \dots (3)$$

and that of per capita income approximately as

$$y_t = y_0(1+\alpha\beta-\rho)^t \qquad \dots (4)$$

where y is per capita income, and ρ the constant rate of growth of population. He used time series data for many countries to work out the α and β coefficients and thus tried to concretise his ideas on the rate of investment and rate of growth of national income in India.

3.5. Simultaneously, by this time as Chairman of the National Income Committee, he has been trying to build up estimates of national income and investment for the country. The problem at this stage was just to get a dimensional picture of the levels of national income and investment. But even this dimensional picture was invaluable because this enabled him to have an idea of the base from which we have to start.

3.6. In the two-sector model[6] (Mahalanobis, 1953b) the entire net output of the economy is supposed to be produced in two sectors, one producing all investment goods and the other all consumer goods. Unlike Marx's scheme, all intermediate goods industries, according to this scheme, are split up between the two sectors such that one could obtain the net output separately for both the sectors, which add up to net national output or national income. The economy, further, is conceived as closed. The new variable introduced is C_t, such that $Y_t = C_t + I_t$, C_t standing for consumer expenditure in the t-th year. It is then assumed that in the t-th year, the net investment is split up between the two sectors in the ratio $\lambda_c : \lambda_i$, $\lambda_c + \lambda_i = 1$. We then have the following two relations of the type given in equation (2):

---

[5] In Mahalanobis' models, the usually accepted definitions of national income, investment and consumer expenditure were systematically used. He has always taken both national income and investment as 'net', i.e., exclusive of allowances for consumption of fixed capital. He is, however, not fully satisfied about this. There could be a more satisfactory concept which will both allow more accurate measurement and be better suited in the context of plan formulation.

[6] A model of exactly this type was developed by Feld'man in the USSR in 1928. A summary of the paper is available in English in Domar's *Essays in the Theory of Economic Growth*. Mahalanobis' work, however, was done completely independently. "It is interesting to note that the impetus of planning initiated almost identical lines of thinking in the two countries, the USSR and India," (Mahalanobis, 1960d).

$$\left.\begin{aligned} C_{t+1} - C_t &= \lambda_i \beta_c I_t \\ I_{t+1} - I_t &= \lambda_i \beta_i I_t \end{aligned}\right\} \qquad \dots \quad (5)$$

where β_c, β_i are reciprocals of sectoral marginal capital output ratios. The above equations lead to

$$Y_t = Y_0 \left[1 + \alpha_0 \frac{\lambda_c \beta_c + \lambda_i \beta_i}{\lambda_i \beta_i} \{(1 + \lambda_i \beta_i)^t - 1\} \right] \qquad \dots \quad (6)$$

where α_0 is I_0 / Y_0, giving the pattern of growth of national income over time.

3.7. From actual observations, it has been found that β_i is always smaller than β_0, and as such the above equation shows a higher rate of growth of national income for smaller λ_i for the first few years. But when one takes a longer time horizon, a larger λ_i systematically gives a higher rate of growth of national income. This model was used as an argument in favour of increasing the allocation of investment in investment goods industries for the Second Five Year Plan. The ratio was of the order of 10 per cent before the plan; and the decision was to raise the ratio to about 33 per cent. Of course, a larger rise would have been desirable for long period ends, but this was not considered to be feasible at the time. The claims of consumer goods industries have to be met and the technological set-up of industries producing capital goods may be such that it is better suited for production of capital goods for the use in consumer goods industries rather than capital goods for capital goods industries. It may take quite some time to introduce drastic change in this set-up. Foreign trade, may, however, come to the aid marginally because one type of capital goods can be exchanged for another type.

3.8. It can be shown from equation (6) that

$$\left| \underset{t \to \infty}{\mathrm{Lt}} \right| \frac{I_t}{Y_t} = \frac{\lambda_i \beta_i}{\lambda_c \beta_c + \lambda_i \beta_i}. \qquad \dots \quad (7)$$

Some limiting values are given below by way of illustration when $\alpha_0 = 6$ per cent, $\beta_i = 0.25$, and $\beta_c = 0.50$:

λ_i	$\frac{\lambda_i \beta_i}{\lambda_c \beta_c + \lambda_i \beta_i}$ per cent
0.1	5.3
0.3	17.7
0.5	33.3

Thus when α_0 is higher than the limiting value, the α_t's $\left(\alpha_t = \frac{I_t}{Y_t}\right)$ will gradually decrease until we reach the limiting value. When, however, it is less than the limiting value, an increasing sequence will emerge. One reason of selecting a figure like one-third for λ_i for the Second Plan is that higher values of λ_i would lead to unusually high rates of investment in the limit. In a planned economy, the effort should be to increase α_t deliberately up to a certain level, and the way to do this is to increase λ_i to the extent possible. "It should be also possible to increase the ratios (β_0, β_i) through deliberate planning. Suppose that in the earlier stage of development, a product of an industry has got practically no market, and therefore is of negligible value. At a later stage of development it may, however, be possible to use this waste product as raw material in some other industry which is established by planning. In an economy which is developing harmoniously, a number of such by-products may gradually become useful. Hence the ratio of the value of net output to the value of raw materials would increase because by-products of the type considered above would continue to be sold at negligible prices. An increase in the ability of the machines or of labour to convert the same amount of raw materials into a larger quantity of final product will produce the same effect, price situation remaining the same. A steady increase in β may be, therefore, a characteristic of harmonious development."[7] (Mahalanobis, 1953b).

3.9. J. B. S. Haldane[8] has made some investigations on the maxima of the function given in equation (6). This work, however, is not of practical interest because of one peculiarity of the function. Generally, with realistic values of parameters $(\beta_0 > \beta_i)$ Y_t at a point of time decreases with increase in λ_t for the first few years, say from year 0 to year T. Then there is a period $T+1$ to $T+\tau$, in which a true maximum exists for each year and can be obtained by following the method given by Haldane. But beyond $T+\tau$ years, the function uniformly increases with increase in λ_t. As T and τ are both not very large, and in particular T is small for realistic parameters, the quest of maxima of the function is not very interesting.

3.10. Mahalanobis' single sector and two sector models are to be distinguished from the models developed by Harrod and Domar, which are frequently given by the same expressions. For example, Domar's simplest model (with our notations) is given by

$$\frac{I}{Y} = \alpha \qquad \dots (8)$$

$$\frac{\frac{dY}{dt}}{I} = \beta \qquad \dots (9)$$

which gives

$$Y_t = Y_0 . e^{\alpha\beta t} \qquad \dots (10)$$

---

[7] Some increase in β has been noticed in manufacturing industries in the USA in the recent past. This may be a result of high and diversified industrialisation. What is possible to achieve by deliberate planning, may also come more automatically when a country reaches an advanced level of industrialisation. Deliberate planning, however, should help in arriving at the position more quickly.

[8] See Haldane, J. B. S. (1955): The maximization of national income. *Sankhyā*, **16**, 1—2.

This is identical with Mahalanobis' $Y_t = Y_0\,(1+\alpha\beta)^t$ except for some higher order terms in the expansion $e^{\alpha\beta}$ contributing very little when α and β are small. While Harrod-Domar models seek to describe how an economy moved in the past, and thus explain a phenomenon, Mahalanobis' models are essentially forward-looking planning models. Mahalanobis' α thus is not to be likened with propensity to save but with the desire a democratically formed Government to invest. Similarly, his income coefficient of investment does not seek to explain investment in terms of growth in income and consumption in the past as is done in accelerator theory but is a proximate estimate of the likely increase in income in future for a certain volume of investment, based frequently on analysis of project data of diverse kinds.[9]

3.11. Mahalanobis (1955a, 1955b) next developed what is popularly known as the four sector model for working out the basic dimensions of the Second Five Year Plan of India. This model is to be regarded as an allocation model and not a growth model in the sense the other two are. This model also is the simplest and is an attempt at using scientific principles in solving a very concrete problem of planning placed before him. It was more or less known that the maximum funds available for net investment during the Second Plan period would be of the order of Rs. 5,500 crores. It was also agreed that it would be necessary to provide for 10 to 12 million new people in the labour force during the period. What Mahalanobis did was to accept these two propositions and add a third, the reasonable rate of increase in national income during the period would be 5 per cent per year. He knew also from the two sector model that a reasonable target for investment in investment goods industries would be roughly one-third of the total investment. He, then, put all these in a simple simultaneous equation system and obtained his solution.

3.12. The sectors distinguished by him and the variables and parameters used are summarised below.

sectors	symbol	new employment	capital needed per worker	total capital needed (3) × (4)	income coefficient of capital	income generated (5) × (6)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. investment goods industries	i	n_i	θ_i	$n_i\theta_i$	β_i	$n_i\theta_i\beta_i$
2. consumer goods industries						
2.1 factory type	1	n_1	θ_1	$n_1\theta_1$	β_1	$n_1\theta_1\beta_1$
2.2 hand type including agriculture	2	n_2	θ_2	$n_2\theta_2$	β_2	$n_2\theta_2\beta_2$
3. services of all types	3	n_3	θ_3	$n_3\theta_3$	β_3	$n_3\theta_3\beta_3$

If N is the total new employment, A is the aggregate net capital formation and E is the aggregate national income generated over the five year plan period, then it is possible to write

$$\left.\begin{aligned} N &= n_i+n_1+n_2+n_3 \\ A &= n_i\theta_i+n_1\theta_1+n_2\theta_2+n_3\theta_3 \\ E &= n_i\theta_i\beta_i+n_1\theta_1\beta_1+n_2\theta_2\beta_2+n_3\theta_3\beta_3 \end{aligned}\right\} . \qquad \dots (11)$$

---

[9] See Ott, A. E. (1958): The relation between the accelerator and capital output ratio, *Review of Economic Studies*, 25(3), No. 68, for a consideration of this topic.

Since, N, A and E are taken to be known, and as it was agreed upon that one-third of the total investment will be in the (i) sector, it is possible to solve this simultaneous equation system. Thus, realistic estimates were worked out of the parameters β and θ on the basis of all available evidence, and the equation system was solved giving the broad sectoral allocation of employment, capital investment and increment in national income. The values of parameters and the solution are given below.

sector symbols	values of parameters used		solution		
	β	θ(Rs.)	employment (N) 10^6	investment (A) 10^9 Rs.	income (E) 10^9 Rs.
(1)	(2)	(3)	(4)	(5)	(6)
i	0.20	20,000	0.9	18.5	3.7
1	0.35	8,750	1.1	9.8	3.4
2	1.25	2,500	4.7	11.8	14.7
3	0.45	2,750	4.3	16.0	7.2
total (given)			11.0	56.1	29.0

After getting the solution, the small scale industries (including agriculture) was split up between two sub-sectors, small scale and village industries proper and agriculture on the basis of certain fresh data.[10]

3.13. The broad sectoral aggregates were then split up into detailed targets as given in the Draft Plan Frame. For this naturally various supplementary and ancillary considerations were necessary. A trial set of targets for a sector was chosen at first and details about employment, investment and income were aggregated. Then the set was altered on a trial and error basis until it was consistent with the

---

[10] The coefficients adopted for agriculture and household industries were as follows: $\beta_h = 2$, $\theta_a =$ Rs. 6250 and $\theta_h =$ Rs. 620 when the subscripts h and a respectively stand for the sub-sectors household industries and agriculture. It may be noted that it is not necessary to make an assumption about β_a. The resulting allocation is given below:

sector symbol	investment 10^9 Rs.	increment in	
		income 10^9 Rs.	employment 10^6
a	946	1083	1.58
h	194	387	3.12

solutions in respect of the sector. This procedure, however, was not followed rigorously. But basically care was taken to ensure that the individual targets within a sector as a whole did not need a higher outlay than was given by the solution and also the aggregate income and employment generated tallied with those given in the solution. Since, not all details were available in respect of all the targets, the procedure sketched here really gave the logical basis of working out detailed targets rather than the actual procedure followed.

3.14. Detailed targets for several industries and services were derived from considerations other than those which were used to obtain the basic solution given in the above table. A rapid expansion of steel industry, for example, was considered necessary because of potential advantages of the country in the field in view of the existence of extensive and rich deposits in the country, the intimate link between the general level of productivity and ability to use machines built out of steel, and shortages of steel felt earlier and consequential drain in the foreign exchange. Similarly for the development of aluminium industry, the need of electricity industry and extreme shortage of copper deposits in the country were the basic considerations. In general, the considerations used for a target, whether relating to a physical product or relating to a service, were given by the experts in the field by way of physical possibilities of increasing output in the case of goods and some assessment of needs of the country in the case of services.[11]

3.15. It may be pointed out that the usefulness of models of this type would depend very much on the accuracy with which the parameters of the model could be estimated. The entire work on the Draft Plan Frame had to be completed in about four months, and due to this shortage of time the estimates of the parameters used had to be tentative, in spite of the fact that large technical resources were put on the task. But the estimates were undoubtedly the best possible ones under the circumstances. Considerably better estimates, however, could have been worked out, had there been more time.

3.16. Many criticisms[12] of the use of these models spring from a failure to grasp one of their essential features. The models were not devised as a contribution to theoretical economics. They are entirely operational models with a very limited object in view, described earlier. Mahalanobis has taken care to point out this repeatedly. But in spite of this, some of the critics have taken pains to point out the theoretical inadequacies of the models, which admittedly exist. The set up of the four sector model, it may, however, be noted, is similar to the general set up of the linear programming problems, and Mahalanobis' problem with some modifications can be tackled in more generality by means of linear programming.[13]

---

[11] Mahalanobis (1955a) gave particular emphasis to health services. He states (Mahalanobis, 1955b), "Under existing conditions it may take 60 or 70 years to provide one doctor for every two thousand persons in the rural areas on an average...I have thought it desirable, therefore, to include in the Draft Plan Frame a proposal to bring some health service to every home in the country within a reasonable time, possibly in 10 or 15 years (instead of providing an exclusively high quality service to a very small fraction of the population), by establishing two new cadres of 2-year and 4-year trained health assistants as a first step to national health service throughout the country...One 6-year trained physician would be in charge of a group of 5 or 6 health assistants...." This scheme would enable placement of one health assistant for every 5 villages in 10 or 15 years at a reasonable cost.

[12] See Mitra, A. (1957): A note on the Mahalanobis model, *Economic Weekly*; Tsuru, S. (1957): Some theoretical doubts on the plan-frame, *Economic Weekly*, Annual Number, 1957; Chakrabarty, S. (1957): The Mahalanobis model of development planning, *Arthaniti* 1, No. 1,; Mathur, P. N. (1957): A note on planning in India, *The Indian Economic Journal*, 4, No. 4; and Sen, A. (1958): A note on the Mahalanobis model of sectoral planning, *Arthaniti*, 1, No. 2, for some topical critical observations on the Draft Plan Frame.

3.17. According to Mahalanobis, working out a consistent and feasible plan-frame is necessary for planning in a real sense in any country. But this is only a part of the entire complex of social activities which go under the name of planning. There are several conditions and limiting factors which are to be satisfied before it is possible to give effect to a plan of this type. Production of commodities needed to satisfy demand for consumer goods has to be undertaken in the diffused and unorganised sectors of the economy and achieving targets would call for vast organisational effort. Since a rapid rate of development without inflationary effects would depend on this, the ability to produce required amount of consumer goods would constitute a limiting factor. Capacity to invest in any given year will depend, by and large, on the ability of domestic industries to produce capital goods and basic raw materials of various types. Thus the rate of expansion of basic industries sets down the second limiting factor. Though it is relatively easier to solve the problem of large-scale production, this limiting factor should be kept in mind. One possible reason why it might be difficult to achieve the desired rate of development is the inability to turn out required number of trained technical and scientific personnel. The problem of trained personnel is also connected with development of consumer goods production and various services. Thus while labour shortages are unlikely to create any trouble in India in the near future, trained manpower would be a very serious limiting factor. It is possible to work out the financial counterpart of a plan which is feasible in terms of available physical and manpower resources because appropriate fiscal and financial measures needed for its implementation always exist. But the ability to raise resources in financial terms would involve financial and fiscal technicalities and socially and politically feasible measures, and measures which exist theoretically cannot always be implemented in practice. This may be regarded as the fourth limiting factor. Rigidities of the existing administrative system furnish the next important bottleneck. Actual implementation of a plan of this type would call for decentralisation of administrative and financial powers in public enterprises and institutions. Also it will be necessary to have active cooperation between public and private agencies. In general, the entire tone of public administration has to be improved and oriented towards planning for economic development, and until this is done, any satisfactory and logically consistent plan would remain, at best, a theoretical exercise.[14] Finally, in the initial phases of development when large volumes of capital goods and raw materials are to be imported from abroad, foreign exchange difficulties would furnish a serious limiting factor and import substitution and export promotion will deserve continued attention.

---

[13] See Komiya, R. (1959): A note on Professor Mahalanobis' model of Indian economic planning, *Review of Economics and Statistics*, 41, No. 1, for an attempt at reducing Mahalanobis' problem as one in linear programming.

3.18. In Mahalanobis' words, (1955b) "The logical consistency of the Plan Frame is not a sufficient guarantee of its feasibility in practice. Any one (or two or more in combination) of the ... limiting factors ... can retard progress. However, so far as plan-making is concerned (as distinguished from plan-implementation) all that can be demanded is internal consistency, valid technical reasoning and a correct appreciation of social needs. If the present plan has these merits, then there is only one single issue, namely, whether there is any alternative plan which would eliminate unemployment and poverty more quickly and more effectively; and at the same time, lay foundations for a continuing increase in the level of living in future. If there is no alternative plan which is more satisfactory then the proper policy would be to try to implement the present plan."

4. THE POST PLAN-FRAME PERIOD

4.1. Mahalanobis continued his deliberations on planning even after the Draft Plan Frame and various associated papers were written. His post Draft-Plan Frame writings fall generally into four categories: (1) He has worked on refining some of his arguments given in the Draft Plan Frame and associated papers. For example, his argument for priority to basic industries has been refined, and the question of saving of foreign exchange has been explicitly brought out in some of his recent writings. (2) He thought about the importance of technical manpower and educational problems even when he was writing the Draft Plan Frame. But subsequently, he has probed more deeply into the subject. As a result of this, he is inclined to think that science has a basic role to play in economic development. He has classified types of activities by time taken to start production after the social decision has been taken and tried to explain economic development in a novel way. Development of science, which takes longest time, remains the corner-stone of his explanation. His analysis contains germs of a forward-looking planning theory, and he has made some historical analysis to support his observations. (3) Thirdly,

---

[14] Mahalanobis (1955a) has given considerable thoughts to the problem of improvement of the quality of public services. He has also written quite categorically on this topic :

"Planning on bold lines with a steady expansion of the public sector and advance to a socialistic pattern of economy would require the building up of appropriate administrative machinery of a new type at all levels.

"There must be decentralization, on business-like lines, of the day to day management of public enterprises with large delegation of financial, administrative and executive control to develop initiative and responsibility at the periphery so necessary for efficient conduct of business enterprises.

"Attention must be focussed on the implementation of the Plan — on getting things done at the right time — and rules and procedures must, if necessary, be revised to ensure effective action. Secretariat control of the present type must be replaced to a large extent by control by truly autonomous public corporations set up by Government or through the supply of credit by State Banks working under the general guidance of Government in matters of policy.

"Administrative difficulties inherent in the existing Government machinery are likely to prove the greatest obstacle to efficient planning. To overcome such difficulties, large organisational and even constitutional changes may become necessary. The problem is urgent and requires immediate and serious attention."

Mahalanobis has written on the world perspective of the economic development of the pre-industrial countries, and has taken pains to show that rapid development of under-developed areas of the world is likely to help in promoting world peace. He has made some analysis of international situation for this, but the theoretical outline considered above also came to his aid.[15] (4) Finally, he has written on the specific problem of industrial labour in India and suggested an interesting solution to our eternal labour troubles and unemployment problem.

4.2. *Priority of basic industries: a re-examination.* Let us consider the case of a country which has to import a given amount of foodgrains every year. The quantity to be imported is obtained in relation to the proximate rate of increase of population and projected estimates of domestic production of foodgrains available for consumption. The requisite quantity of foodgrains has to be purchased at the prevailing world price and hence the cost involved will be the product of the quantity, the world price and the number of years for which the purchases have to be made. Provided the commodity concerned is available in the world market in sufficient quantities, the decision about importing may be taken immediately before the period of consumption. In case it is possible to take decisions earlier, at least by one year, the import of foodgrains may be substituted by import of fertilizers. In India, application of one ton of fertilizer can be expected to produce two extra tons of foodgrains. Since world price of a ton of fertilizer is less than that of a ton of foodgrains, there is considerable saving in foreign exchange in the second variant. The arrangement about distribution will entail administrative measures both in the case of importing foodgrains and in the case of importing fertilizers and the costs involved in the two variants may be of the same order.

4.3. If it is possible to take decision, say, five years before the year of consumption, it is possible to build up fertilizer plants to produce the requisite amount of fertilizer every year by importing necessary machinery from abroad. The foreign exchange requirement of this variant will be considerably less than that in the second variant for any sufficiently long period of years. The distributional problems will remain more or less the same in this case. It should, however, be possible to build up the plant; i.e., raw materials etc., should permit economic exploitation, funds for investment should be available, and there should be no shortages of labour including technical manpower. In India, all these conditions are fulfilled for fertilizer production, and hence construction of fertilizer plants with imported machinery is more economical to the country from the foreign exchange point of view than importing fertilizer from abroad. The overall cost also will be smaller in this variant provided we take a sufficiently long time period in consideration. As labour is relatively cheap, raw materials are available in the country and trade margins and transport costs are likely to be lower, the price of domestically produced fertilizers should be lower than the price of the imported fertilizer.

---

[15] Mahalanobis has covered these topics in (1958; 1959b; 1959c; 1959d; 1959f; 1960a; 1960c; 1961a; and 1961b). The same topic has sometimes been considered in more than one paper, and as such we refrain from giving detailed references here.

4.4. The last alternative is to build up the capacity to produce machines in the country and then produce fertilizer plants of requisite capacity with domestically produced machines. This will require decisions to be taken at least some ten or fifteen years before the year from which additional consumption is needed because machine building industry has to be developed first and then the fertilizer plant has to be set up. Also since, an integrated plant for building fertilizer plants may not be economical, it may be necessary to build up a complex of machine building industries producing not only fertilizer plants but also other plants. If this is done, the foreign exchange cost (apportioned) will be the minimum for any reasonable period of time. The procedure, however, will require large aggregate investment outlay and a training programme for engineers and technicians. Also, raw materials etc., should be available for developing a machine building industry, a condition which is satisfied in India. The allocated aggregate cost in India for building up a fertilizer plant is likely to be smaller than when the plant is built with imported machine for reasons outlined in the last paragraph.

4.5. This is the central theme of Mahalanobis' (1958) arguments in "Science and national planning." To sum up, earlier a decision is taken, it is possible to choose a variant which is more economical to the country in the long run, both from saving of foreign exchange and also for other reasons. Of course, choice of each variant is associated with a particular complex of associated tasks not all of which are easy. But supposing decisions are taken and the tasks are squarely faced and fulfilled reasonably well, then it is always better to take decisions as early as possible and in general, choose a variant in which the basic capital goods are produced in the country. In fact, this is a quintessence of the philosophy of planning, and is probably profoundly true, provided there is planning in a real sense. Planning in this context means that democratic decisions on plan targets and plan measures are taken and adhered to strictly. Decisions, however, are not to be taken about a period which is so remote that it is impossible to think concretely about the period. Some risks are involved in taking early decisions; but such risks should be calculated ones. No country in the world is planning to do without electrical energy with a hope that controlled thermo-nuclear power would provide all the energy we need one day.

4.6. *Scientific research, technical manpower and education.* Setting up complexes of heavy industries, as we have already noted, require decisions, among others, on the training programme for engineers and other technical workers to man the industries. For this, it is necessary to set up technical colleges and institutions and train up teachers for such institutions. While setting up basic industrial complexes could be accomplished in ten or fifteen years when a firm decision is taken and there is no dearth of technical manpower it is necessary to take decisions fifteen or twenty years in advance, if the desire is to man the complexes by domestic personnel. Since, there is little chance of finding foreign personnel for large industrial complexes, the only practical way is to think of the personnel also in the very beginning, that is, take decisions fifteen or twenty years in advance rather than ten or fifteen years in advance. However, it is not possible to develop such teaching programmes in a country in which scientific research has not developed, and as such it is necessary to think of development of scientific research in the country for economic development.

4.7. A consideration of this type led Mahalanobis to formulate what he has aptly described as the four-fold or five-fold logic of economic development. The following paragraphs taken from "The need of scientific and technical manpower for economic development" (Mahalanobis, 1959d) give a concise exposition of the principle.

4.8. "We . . . have to think of four levels. First, to increase the supply of consumer goods which, so to say, is at the top or the first level. To do this, we must expand the production of capital goods; this is the second level. Both of these will require a larger and larger supply of engineers, technologists and technical personnel; this is the third level. Engineering and technological development would call for an increasing volume of applied research. But applied research requires a sound foundation of basic research. We must have an increasing supply of research scientists of ability. Unfortunately, their number is small in every country. We must try to make the best use of all whom we can discover. This is the fourth level.

4.9. "Now consider the factor of time. We can set up factories for consumer goods very quickly; in a year or two, if we use imported machinery. To develop the production of capital goods would take more time, from 5 to 10 years at least. To secure an adequate supply of engineering and technical personnel would require still more time. And finally, we must have enough scientists of ability for both applied and basic research which would take at least a generation. This is the four-fold logic of economic development.

4.10. "How to attract and hold a sufficient number of able persons in science and technology is then the crucial problem of national development. This can be done only through a proper social appreciation of science and scientists, which is the fifth and deepest level of the problem."[16]

4.11. It will be seen that Mahalanobis considers scientific advance as the primary pre-requisite of a sustained economic development and hence it will be useful to dilate further on this point. Economic progress entails a steady growth of national product, and this is achieved by transforming various material given by nature into capital goods and consumer goods. This process is technological in nature and engineers and technologists have a vital role to play in the process. Also natural resources vary widely from country to country and proper and efficient utilisation of national resources given in a country involves sustained technological research. Such research could only prosper in an atmosphere in which science in general is flourishing and the stock of scientific knowledge is advancing. Thus promotion of pure research automatically creates a condition in which more utilitarian research could prosper; it is for this reason that one should think of a close link between the economic development and pure research. Organisation and promotion of scientific research in its most comprehensive sense is thus an essential requirement for rapid economic development.

---

[16] It has also been pointed out elsewhere (Mahalanobis, 1961a) that higher priority should be given to tiers or levels which are more slowly maturing. Thus talking about the fourth level, he states, "This is the most slowly maturing sector, and must, therefore, be given the highest priority."

4.12. As research is closely linked with economic development, Mahalanobis made a pioneering attempt to quantify the extent of research in a country and notice its relation with the economic development. He has studied (Mahalanobis, 1958) the indicators of economic development and the aggregate research outlay of five countries, USA, USSR, UK, China and India. He has noticed that for rich countries with high per capita national income and large per capita consumption of energy, the aggregate expenditure on research is of the order of two per cent of the national income. In poor countries like China and India, the share is extremely low being about 0.4 per cent in China and less than 0.1 per cent in India. Thus, according to him, for under developed countries, it is essential to raise this ratio substantially in order that sustained economic growth is possible. This line of research initiated by him has subsequently been taken up by others and an important field of study has been opened.

4.13. Promotion of scientific research, however, is not an easy task in a country like India because this involves social acceptance of newer values. "The value of science to society lies in its unorthodoxy and ability to challenge accepted concepts and theories" (Mahalanobis, 1960c). In most human activities other than scientific pursuit, best exemplified by administration, a subordinate has to obey his superior. This is also true in judicial and religious matters. Decision of a higher court overrides that of a lower court. Higher the level of a priest, the more is his authority and prestige. In fact, prior to the scientific revolution brought about during last four centuries, most human activities were organised in a hierarchical system in which the view of a person in a higher position in the hierarchy would be accepted against that of a person in a lower position, merely by virtue of the position. There could be no appeal outside the hierarchy; the only appeal lay with persons still higher up in the hierarchy. It will be seen that for many human activities, this principle remains essential even today. For example, it is not possible to have an efficient army without accepting these tenets. The scientific revolution brought about a new concept of objective validity which has its foundation in nature itself and which cannot be changed or altered by persons having a higher status. The modern age is essentially based on this notion of validity. "A linear order of ranking and of levels of authority is indispensible in administration. The opinion or decision of a superior officer must prevail. This is in complete contrast to decision making in science in which the ideal must be complete equality of status of all scientific workers, as all decisions must be based on reason and not on authority" (Mahalanobis, 1960c). It has to be underlined that this principle is applicable only to science among numerous human pursuits. The objective nature of science requires promotion of democratic principles for its progress.

4.14. Even when there is social acceptance of this principle, there is another important obstacle in the way. For sound organisation of science, it is necessary to have an objective method of assessment of scientific workers and their work. A necessary condition for this is the existence of basic scientific outlook and a free atmosphere of criticism. Appropriate traditions of assessment and criticism developed in many western countries during last four hundred years. Many important scientific institutions in these countries were completely outside government influence; this enabled them to build up a standard of their own and also to help government by offering independent criticism of government scientific policies. In the Soviet Union and other socialist countries, the Academy of Sciences and their associated institutes were deliberately developed as independent institutions outside the sphere of government control. This enabled them to develop adequate standards quickly. The position in India is still unsatisfactory in this respect; but this remains a very important problem which has to be solved before we can hope to be able to progress rapidly towards our economic goals.

4.15. We have noticed earlier that the training of technical manpower is intimately linked with long range planning of a country. But the basic problems enumerated above has to be solved before a real headway could be made towards training of technical manpower including research personnel. Certain educational reforms are needed in India in order to help in the process. It was suggested by him that no deserving student should be deprived of educational opportunities because of poverty. This will require free education at all levels. Since it will be difficult to do this at the present stage of our development, the higher education should at least be made free during the Third Five Year Plan period. This will entail abolition of all fees at the university level and devising a system of scholarship which will sustain all deserving students during the period of their training. This will also involve selection of students for higher education entirely on merit.

4.16. "Inequalities in medical care result in a larger number of deaths among the poorer people. Inequalities in educational opportunities have a deeper and more pervasive effect. It is true that some extremely able students succeed in securing scholarships upto the highest level but their number is very small. By and large, it is the rich people who have the opportunity of giving their children the type of education required for posts of influence and responsibility in the country. By and large, those having such training are selected for posts of responsibility on the strength of their higher educational qualifications. In this way the power and privileges of a small group of people at the top tend to be not only preserved but strengthened. In both public services and organised private enterprises practically all posts of influence and power are held by persons belonging to the same small privileged class. This has created an influential group of people who naturally desire to maintain their privileged position and power. During the British period many of the influential people were not enthusiastic about the political change because they were afraid of losing their own privileges. In the same way, it is not surprising that there are people in India today who are not enthusiastic about a rapid economic progress out of a similar fear of losing their privileges and power. It is necessary to remove barriers to educational opportunities to overcome such difficulties. This is the only way in which a sound foundation can be laid for democracy and socialism in India" (Mahalanobis, 1960a). I have quoted this in full because apart from indicating the reasons for the educational reforms suggested by him, the paragraph underlines his deeply human and democratic approach towards all problems of planning.

4.17. As I have pointed out earlier, what Mahalanobis has done is to give an indication of a theory, and not any worked-out theory of economic development. In his mind, scientific advancement has a very important and basic role to play in the technological innovations which led to rapid growth of output and productivity of UK and some other nations of the world. This led to a large increase in the level of living of the Western countries; thousands of years of stagnation in the per capita income was replaced by new trends of rapid economic growth over the last century and a half in the Western countries. This gave the Western countries undisputed military supremacy which enabled them to dominate the rest of the world into either direct colonial rule or into conditions of economic and political subjugation.

4.18. Mahalanobis has not directly brought in the question of cause and effect into the picture. He has stressed more on simultaneity. In the beginning of the industrial revolution in the 18th century, revolutions in the technologies of spinning and weaving first, then in iron and steel and then in electricity in the 19th century made possible the phenomenal growth in output and productivity. Simultaneously there was emergence of the bourgeoisie and the capitalist class with their aspiration towards social and political power. He, however, ventures to suggest that probably the former stimulated the latter. Also, the spread of scientific outlook probably "prepared the ground for the age of reason and the French revolution towards the end of the 18th century; this led to the growth of nationalism, in its modern sense, in Europe in the 19th century" (1961a). The industrial revolution led to the use of machinery and thus replacement of animal and human power by steam and electrical power. The link between technological development and scientific advance became very close in the 19th century. Mahalanobis specifically admits of the interacting nature of the industrial and technological advances, without pointing out which is prior. He states, for example "industrial development is being sometimes stimulated by a new scientific discovery, or scientific discovery is being stimulated by industrial needs" (1961a).

4.19. To sum up, while the role of innovations in economic development has been stressed by eminent theorists like Schumpeter, Mahalanobis has pointed out the interrelations between scientific advances and technological innovations. He has also brought out clearly the role of decision making at appropriate time. However, one should point out once again that Mahalanobis' analysis is not primarily intended to explain what has happened in the past. It is entirely forward looking like his planning models and gives a theoretical frame work to the Indian Planners, in order to help them in their work. It also follows that these considerations apply only in a planned economy.

4.20. *Industrialization of undeveloped countries and world peace.* Subsequent to the October revolution in 1917, USSR adopted a policy of planned economic development and by and large achieved phenomenal success. Other socialist countries followed the suit after the Second World War. In general, the socialist countries succeeded in achieving a larger rate of development in the recent past than the capitalist countries in the post industrial revolution period, thus showing the advantages of planning as a method of economic development. In order to have planned progress, USSR and other socialist countries deliberately adopted a policy of encouraging science and technology by creating a condition in which these pursuits give great prestige. They also adopted long range programmes for the development of technical manpower. These measures support Mahalanobis' analysis of the situation as given earlier. As a result of this, the monopoly of Western countries as the only world power has gone, and socialist countries, as a block, has emerged as a second power to reckon with. This has led to world tension and a state of cold war. Mahalanobis, while thinking about the question of economic development in the context of the world as a whole, was naturally confronted with this situation. He has, however, introduced a fresh point. One basic, and perhaps the most potent reason of the world tension, is the existence of underdeveloped countries of the world. Both Western and Eastern powers are interested in increasing their spheres of influence among these countries, and various diplomatic methods and aids of diverse types are being utilised for this purpose. This is leading to clashes between the Western and Eastern powers, and hence helping to prolong the world tension. In many of the underdeveloped countries, the ruling group is not very stable and is liable to change at a slight provocation. Pressures and counter pressures from Western and Eastern powers, therefore, naturally develop inflammatory situations in these countries, and many incidents in the underdeveloped countries in the last decade may partly be explained in this way.

4.21. Mahalanobis has emphasized that the surest way to avoid tension on this count is to ensure rapid economic development in the underdeveloped countries of the world. This will lead to stable leadership in these countries. Also if the democratic opinion is satisfied about the progress, the condition within the countries, as a whole, will be stable. This will enable them to react more rationally to political and other pressures from outside.[17] Some of the developed countries in either block may be interested in reduction in world tension, and they may help in securing economic development of underdeveloped countries. But it has to be realised that rapid growth will entail planning of the type which he has in his mind, i.e., with emphasis on scientific research, technical manpower and basic industries, and the advanced countries has to realise this in order to be able to help effectively. This is the best possible way to secure lasting peace in the world.

4.22. It has been shown by Mahalanobis and some others that a small percentage of world defence outlay is spent on uplifting the conditions of the poorer countries of the world; the results will be spectacular. This is easy to see when it is noticed that the world defence outlay per year is of the order of three or four times the entire national income of India or more than double that of China. Thus, if there is agreement between Eastern and Western powers, then one way of securing lasting world peace would be to reduce outlay on arms and ammunitions and increase the outlay on development of poorer regions of the world. Mahalanobis has attempted not only to elaborate this view but also to propagate this view before a wide forum.

---

[17] See also (Mahalanobis, 1959b), "A very small group of families or persons have the largest share of wealth, income and political and economic influence. In fact, the greater the lack of economic development the fewer would be the number of persons who have the effective power of making political and economic decisions. This makes it possible for a foreign power to exert pressure on a small group of powerful persons to give concessions in favour of the foreign power. Such arrangements, because they depend on the will of only a small group of persons, are necessarily subject to violent changes from time to time. Relations between foreign powers and underdeveloped countries are, therefore, basically unstable.

4.23. *Labour problems and unemployment.* Mahalanobis (1959e; 1960b) has explicitly considered labour problems with particular reference to unemployment. The notion of unemployment is admitted to be different in India in comparison with advanced countries mainly because of the existence of a large body of self employed population. "As they do not have any jobs, they cannot lose jobs, and cannot therefore be unemployed in the sense of the industrialised countries" (Mahalanobis, 1959c). On the other hand, it is very well known that manpower is grossly underutilised in the country, and both employees and self employed workers put in much less work than they could. An urgent problem posed in the Second Five Year Plan was, therefore, to make fuller utilization of the available manpower to the tune of equivalent of 11 million regular jobs. This target has not been realized, and the problem of fuller utilization of manpower still looms large before the country.

4.24. The condition of living of the working class has not probably improved materially in comparison with the pre-war period and labour productivity has probably remained stagnant, thus introducing a vicious circle. There is a general disinclination for hard work, and overstaffing is rampant in enterprises, which is sometimes justified on grounds of unemployment and underemployment. Mahalanobis is conscious of this stage of affairs, and he pleads for improvement in efficiency, productivity and remunerations simultaneously. "In an underdeveloped country like India, there is an urgent need to make the whole social organisation more alert and efficient" (Mahalanobis, 1960b).

4.25. Mahalanobis has put forward an interesting idea of labour reserve service as a solution of labour problems in industry and other organised enterprises. The service is to be maintained by the government out of funds provided by enterprises and workers themselves supplemented by general revenue. The reserve will be utilized for productive work and hence it will be able to finance a part of its cost out of sale proceeds of goods produced. When a particular worker is thrown out of job by an enterprise for reasons of efficiency or other allied reasons, the enterprise will have to make a payment to the labour reserve service which is considerably lower than his actual remuneration when he was in the pay-roll. The enterprise, however, would think of dismissing a person only when it would consider the benefits accruing is more than the direct contribution to be made. The workers joining the service would draw a pay from the service which is lower than their original pay but is reasonably sufficient for sustenance. They would also receive training while working with the service. Eventually, with industrial expansion, persons in the reserve will continually be absorbed in other enterprises, and the training will be oriented towards this. The labour reserve service will thus be a buffer against unemployment and would serve as an unemployment insurance scheme for the workers concerned. Since, enterprises will be able to dismiss a worker on grounds of economy and efficiency for a small contribution, the chronic labour relations problem will be avoided and the enterprises will be able to work for highest attainable efficiency and productivity. The scheme thus will help both the workers and the entrepreneurs. Since industrial efficiency remains a serious problem with us, and improvement in productivity is essential for our industrial progress, particularly for increasing our foreign earnings, it is necessary to adopt a policy in the field of labour relations in which these goals can be attained without affecting the level of living of the workers. Mahalanobis' scheme attempts just to do this and deserves serious attention.

5. Concluding remarks

5.1. Mahalanobis' contributions in the field of planning and allied subjects have been wide and varied. He has helped in setting up a sound statistical system in the country; he has prepared a concrete framework for the Second Five Year Plan and he is actively associated with the preparation of a long period perspective of economic development of the country. His role in laying down a sound foundation for statistical reporting is of importance because realistic planning is not possible in the absence of reasonably accurate data. Some of his ideas incorporated in the Draft Plan Frame for the Second Five Year Plan have been accepted by the country and have become a part and parcel of our general thinking on planning. Even his mathematical scaffoldings, which he did not consider as very important, contain ideas which may prove useful in future. The various limiting factors enumerated by him in his Plan Frame narrowed down the field of our future struggle for planned development. His ideas on long range planning with emphasis on scientific research, training of technical manpower and basic industries are gradually gaining acceptance and are likely to help us in achieving sustained economic development.

5.2. I think it will be true to say that Mahalanobis has helped materially in giving a concrete shape to Jawaharlal Nehru's general ideas on socialistic planning. Planning in India would have been more conservative but for his influence and this is all that is claimed as social influence of his work. His influence in theoretical and methodological fields is wide. But we can possibly sum up his contributions by pointing out that he stands for a rational and scientific approach towards planning in India in an atmosphere in which intuitive approach and rule of thumb still play a large role. In planning as well as in his multifarious other activities, he has remained true to science.

5.3. I may perhaps conclude this essay by pointing out one aspect of Mahalanobis' scientific approach. Though he is capable of reaching theoretical heights, he is essentially an applied scientist and he would always like to see that his scientific ideas and theoretical formulations are tried in practice. While working in the field of planning, he was not satisfied with preparing a logically consistent and theoretically valid Draft Plan Frame. He was also interested in having the Draft Plan Frame tried in practice. An applied scientist, his theories have to be experimented upon and supported by empirical observations. But when the field of study happened to be planning for national development, this meant a tremendous amount of effort involving meeting of hundreds of people, engineers and technicians, administrators and politicians. And it is because of this that he had to take upon himself a role in the official planning hierarchy of the country.

A LIST OF PAPERS AND ADDRESSES BY P. C. MAHALANOBIS ON PLANNING

(1950): Why statistics? Presidential address, Indian Science Congress. *Sankhyā*, 10, 195.

(1952): National income, investment and national development. Lecture at the National Institute of Sciences of India. *Talks on Planning*, Indian Statistical Institute, 1961.

(1953a): Studies relating to planning for national development. *Talks on Planning*, Indian Statistical Institute, 1961.

(1953b): Some observations on the process of growth of national income. *Sankhyā*, 12, 307–312.

(1955a): Recommendations for the formulation of the Second Five Year Plan. *Talks on Planning*, Indian Statistical Institute, 1961.

(1955b): The approach of operational research to planning in India. *Sankhyā*, 16, Parts 1 and 2, 3–62.

(1955c): Approach to planning in India. Radio talk, *Talks on Planning*, Indian Statistical Institute, 1961.

(1956): Statistics must have a purpose. Presidential address, Third Pakistan Statistical Conference.

(1958): Science and national planning. Address at the National Institute of Sciences of India. *Sankhyā*, 20, 69–194. *Talks on Planning*, Indian Statistical Institute, 1961.

(1959a): Heralding a new epoch. *A Study of Nehru*. Also published in *Talks on Planning*, Indian Statistical Institute, 1961.

(1959b): Industrialization of underdeveloped countries — a means to peace. Third Pugwash Conference, Kitzbuhel, Vienna, *Bulletin of the Atomic Scientists*, 15, No. 1, and *Sankhyā* 22, Parts 1 and 2, 1960, *Talks on Planning*, Indian Statistical Institute, 1961.

(1959c): Study of the problem of industrialization in the underdeveloped countries. *Contemporary East*, Moscow, *Talks on Planning*, Indian Statistical Institute, 1961.

(1959d): The need of scientific and technical manpower for economic development. Radio talk, *Talks on Planning*, Indian Statistical Institute, 1961.

(1959e): Unemployment. Chairman's address, All India Labour Economic Conference. *Indian Journal of Labour Economics*, 2, No. 1, *Talks on Planning*, Indian Statistical Institute, 1961.

(1959f): Recent developments in the organization of science in India. Indian Statistical Institute, 1961.

(1960a): Next steps in Planning. Address at the National Institute of Sciences of India. *Sankhyā* 22, Parts 1 and 2, 143–172, *Talks on Planning*, Indian Statistical Institute, 1961.

(1960b): Labour problems in a mixed economy. Presidential address at the Third All-India Labour Economic Conference. *Indian Journal of Labour Economics*, 3, No. 1, *Talks on Planning*, Indian Statistical Institute, 1961.

(1960c): A note on problems of scientific personnel, Chairman's note placed before the Scientific Personnel Committee, Planning Commission, Government of India.

(1960d): Operational research models used for planning in India (with M. Mukherjee). Second International Conference on Operational Research, 1960.

(1961a): Role of science in economic and national development. Lecture at the University of Sofia. *Ind. Journ. Pub. Adm.*, 8 (1962) No. 2, 153–160.

(1961b): Statistics for economic development. *Journal of the Operations Research Society of Japan*, 3 No. 3, 97–112

A PHILOSOPHICAL VIEW OF CIVILIZATION

By RADHAKAMAL MUKERJEE
University of Lucknow, India

The meaning and values of civilization are not easy to define. Difficulties arise due to the elusiveness of the triangular interchange between Man, Values and Nature in terms of which civilization can properly be interpreted. There is something free and unpredictable in these triple components of civilization. Civilization comprises the dynamic interchange or "transaction" between these three open entities —Human Nature, Values and Environment, physical and ideal— and can be understood only as their 'emergence' or 'transcendence.' This defies the physical science methods of analysis and treatment almost invariably adopted in the current analysis of civilization. Civilization is fundamentally open, emergent or transcendent, issuing forth into totalities, continuities and potentialities and cannot be causally explained. Nor can any scientific sociology, politics and economics usher in the unity of mankind and create the world community—the goal and aspiration of modern civilization—without an open, expansive value-orientation that only a philosophy of human nature and values can envision and formulate.

Our scientific and technological age views civilization in East or West largely in the context of man's control and manipulation of Nature. The furniture and appurtenances of civilized living, the scaffolding and the frame-work rather than the mansion itself loom larger in the picture. The mansion of civilization is, however, represented by the insights, appreciations, values and aspirations of mankind. It is these enduring experiences of the race which embody the essence of civilization and should obtain primary consideration in any adequate treatment.

The empiricism of modern physical sciences has made Nature qualityless, meaningless and purposeless. It not only injects a corroding sense of doubt and nihilism into human civilization, but also has a desiccating effect on sociological thinking and imagination in every field by emptying them of all considerations of value and purpose in the name of "scientific objectivity." As the philosophy of science and technology makes a tour round the globe, the meanings, appreciations and values that civilization actually experiences in its orientation to cosmos and to existence as a whole are discounted as metaphysical and religious—irrelevant for modern secular culture. The latter seeks a physicalist interpretation of all arts, functions and experiences of human living. This weakens or destroys historic myths, rituals and symbols by depleting them of their constructive cultural meanings and values, and thins out the aesthetic and spiritual resources of personality and civilization.

An exaggeration of the "techniways" and an economic interpretation of civilization have paved the way towards Marxist historical materialism. This is based on the affirmation of a rigid dialectic of productive forces and relations now deployed to explain everything in civilization from government, law and status system to knowledge, art, religion and morality. Dialectical materialism has today its exponents among Oriental countries and peoples far distant from its place of origin.

If the lop-sided Marxist interpretation of civilization disregards the reality of human values and ideal ends of society, the sociological interpretation is no whit better. Contemporary psychology offers in many respects an incomplete and implausible view of human nature and strivings and the values and possibilities of the human individual due to the pressure of physicalist and biological assumptions. Yet it audaciously pushes its claims as a complete science of human behaviour. Contemporary social science in dealing with society and civilization is deflected by the archaic premises and presuppositions of both biology and psychology. It exaggerates man's drive and tension-reduction and opportunistic adjustment at the bio-social dimension, and neglects value-seeking and ideal value-experience. It altogether overlooks the emergent multidimensional character of his adjustment and values and his native oneness with life, society and cosmos.

Civilization establishes an enduring harmony and wholeness of meanings, values and aspirations necessary for both the unity of human self and the integration of life-community-and-cosmos as a whole towards which the evolution of life, mind and society leads. The same evolutionary process which in the inner life of the personality builds up the whole or universal person or real Being—we encounter him in the privileged eras or peaks of every civilization—establishes also a world community that passes on into an unlimited cosmos. An appreciation of the growth and maturation of the individual from this perspective makes the adventure towards a world civilization and solidarity of mankind—man's unique ecological and ethical responsibility. The total movement of man, both as individual and as species, converges towards universality and transcendence. His biological, moral and metaphysical aims coincide.

For this interdependence and convergence in man's ethical evolution and civilization, humanistic metaphysics or religion is the major impulsion. In the growing world age we cannot have several religions, but a world humanistic religion or rather a world humanistic philosophy of religion, grounded in the mutual interchange, agreement and revaluation of spiritual symbols, truths and values in man's mystical contemplation irrespective of his religion and culture.

The ethical problems of contemporary civilization have arisen due to the lack of unity and balance between bio-social and religious, between instrumental and intrinsic, between sensate and transcendent values. The unbalance and conflict are underlined by the outlook and evaluation of contemporary science and philosophy and the rise of a world-wide secular culture to the detriment and lapse of the intrinsic and ultimate values. The integration and realization of the intrinsicalities of life can alone confer a new unity of fulfilment on the emerging world mind and personality of the individual in different cultures, build up a universalist ethics, and promote the unity of civilization through a single movement and struggle of revaluation which for the first time in human history has become in every sense global. Historically, civilizations, Eastern and Western, have always met. In the present world-age civilizations, Eastern and Western, are interchanging and interpenetrating in a global revaluation of value systems. On the moral dimension civilization is undivided and indivisible.

Only a blend of the philosophical-normative with the scientific-empirical methods and principles can properly interpret the dynamic interchange between Person-Values-and Cosmos in their progressive convergence in the coming world civilization. The natural science methods and principles in the study of human civilization have yielded a strangely shallow and sophisticated theory of ethical and cultural relativism now ruling in social science, viz., that it is impossible to conceive of absolute truths and norms independently of the values and the worth of man, unrelated to his social and historical context. There are genuinely universal and inalienable elements of human worth and dignity in all cultures—absolute truth, love and goodness that transcend any particular societies and peoples. Only the philosophical method can combat many harmful and disruptive Western doctrines of relativistic historical ideologies and reveal a common teleological world of values that embodies itself in the community of mankind-and-cosmos and the moral and spiritual unity of civilization. "One World" needs today a new discipline of *Meta-civilization.* This will be grounded in *Evolutionary Transcendence* rather than in Evolutionary Naturalism, and will adopt philosophical rather than scientific procedures for the analysis of the unities, interchanges and transcendences of each of the triad, Person-Values-and Cosmos in the dynamics of civilization.

Biologically, civilization implies man's transcendence of the mechanisms of natural selection of the human animal and transformation of his innate nature and fulfilment of his new evolutionary possibilities enlarging through his heritage of symbols and values. Psychologically, civilization implies his self-surrender and self-transcendence along with self-direction and self-actualization and achievement of unity, poise and freedom in his inner life through realization of the intrinsic and ultimate values, transcending the instrumental and proximate goals and limited boundaries of the immediate biological and social situation. Spiritually, civilization is the realization of man's essential Being as the One, Real and Whole as he achieves harmony and filiation with the mankind-and-cosmos total. This provides the ontological basis of ideal values experienced by man as natural and positive norms, laws and imperatives in civilization. In this process civilization becomes also invested with the unity, universality and immortality of the Absolute Being.

Each phase or stage in man's multidimensional civilized living is linked with the other. This process that balances and integrates various orders and dimensions into a harmonious and convergent whole may be called *Transhumanism :* man, both as individual and species, surpassing his ecology, psychology and sociology, and deliberately choosing in the adventure of civilization the dimension not of actualities, but of potentialities of, and for his nature, relationships and values.

Transhumanism stemming from the unattainable needs, values and visions of the human mind is the core of the quest of civilization and its aspiration. It reveals simultaneously the finitude and imperfection of man and civilization as well as their aspiration for unrealizable goals and values, the incompleteness of harmony between man and society and cosmos as well as the unflagging striving for fresh communion and solidarity. It is the intuition and feeling of Transhumanism which ever lead man to yet more complete Truth, Beauty and Love and civilization, to new adventures of Goodness, Peace and Harmony.

Civilizations, as both Burckhardt and Toynbee observe, do not die of mere senility or weakness. It is the lapse of ideal values which represents the failure of a civilization and causes its ultimate destruction. Modern mechanical-industiral civilization exhibits symptoms of decline because of its false value-orientation, the pride of power and possession and worship of material goods and comforts—the instrumentalities of life. Affluent, mighty nations now seek to build the world community on the basis of socio-economic systems, traditions and values that belong to a particular epoch of a particular civilization—the contingent and finite phases and circumstances in world history. Their strategies and techniques lead to exploitation, oppression and slavery. These are the vices of Imperialism. Or they raise false doctrines of class struggle and economic unification of mankind to the status of absolutes. Their strategies and techniques bring about social antagonism, violence and revolution. These are the vices of Communism. Both the world strategies, competing, overlapping and intersecting, bring about human coercion and suffering on a global scale. Into this dismal picture enter the dire possibilities, long hidden, of the mutual extermination of the colossal rival powers, each striving for world hegemony, through the use of atomic science and technology and the rebarbarisation, if not the extinction, of mankind. Science and technology, practical skill and knowledge without love, compassion and reverence serve merely to augment the wickedness and suffering of man and civilization. Through an upsurge of subhuman impulses and feelings the noblest in human life is fashioned into an evil that did not exist before.

Wherever civilization thwarts or denies open, ideal values and the freedom and universality of the real self and seeks to complete human history without regard to the finite character of the empirical individual self and of the specific collective adventure, it becomes distorted and wicked and is ultimately destroyed. Transhumanism is completely opposed to the perverse and disastrous idolatry or moral and intellectual worship of particularity and immediacy instead of universality and eternity. Every civilization, before its historical decay and destruction, reveals implicit, dangerous elements of dogmatism and intolerance in the intellectual dimension, of infatuation with pleasure, possession and power in the moral dimension, and of abuse of knowledge, faith and creativity in the spiritual dimension. To the extent a civilization accepts the possibilities of Transhumanism instead of idolatrous worship, it renews individual and collective freedom, creativity and adventure, and extends its life indefinitely. By and large history is the record of the tragedies of civilization. But these need not necessarily imply a tragedy for mankind as a whole. For some portion of mankind can and does renew itself through the pattern of ideas and values of Transhumanism. The spiritual visions and virtues of Transhumanism hold the keys of mankind's destiny.

Transhumanism is a forward-oriented human evolutionary necessity, ever spurring creativity and adventure and envisaging and seeking human perfection ever excluded from the completeness of physical attainment. It is constantly stimulated by man's ontological insight and imagination, momentous in their creation and co-ordination of values and ideals for human advance; while its impulsion as a world teleological force, derived from the identification of subject and object, person and Being is akin to that of a world faith. Neither Ethical Humanism from Comte to Bertrand Russel, which, though humane and liberal, is principally confined to the intellectual dimension, and hence cannot bear the burden of the world's anxiety and despair, nor Naturalistic Humanism from Marx to Freud which has become an anthropocentric, myopic and sterile creed, but Transhumanism, aiming at self-transcendence through man's creative and harmonious living in the unbounded society of the cosmos, can offer an effective guidance to civilization today.

Transhumanism can silence many acute and ancient dualities or oppositions between natural and supernatural, flesh and spirit, self-love and self-transcendence, freedom and cosmic order. Those flagrant dissociations are particularly true of the West where these have been reinforced by Cartesian and Kantian rationalism, social Darwinism and Freudian ego-psychology, causing pervasive frustrations and blocking human adventure in every dimension of life. Transhumanism can alone properly direct the education, strivings and aspirations of mankind towards its hidden potentialities, undeterred by the contemporary tragedy in human nature and social experience. It may invest itself with a startling certitude and Messianic fervour, rooted in the rhythms and harmonies of the cosmos that belong to Being or Reality, and that show a strange continuity with man's concrete perceptual and thinking processes that science only recently has begun to recognise.

Civilization at its culmination has a sense of crampedness, discord and alienation within its boundaries and a zest for extension into a boundless community of earth and cosmos. Man reaches his highest when he thinks, feels and lives beyond humanity, when infinitude and eternity exist in his spiritual dialectic at each Here and Now

a World in a Grain of Sand,
And a Heaven in a Wild Flower

A folk-song from Bengal says: "The One is finite and unbounded at the same time. In all ages you may count the moments, in every moment find the eternity. The drop is in the ocean, the ocean in the drop." "The centre is everywhere. Bent is the path of eternity," observes Nietzsche.

A CRITICAL APPRAISAL OF THE CONCEPT AND THEORY OF UNDEREMPLOYMENT[1]

By GUNNAR MYRDAL
University of Stockholm

SUMMARY

The concept and the theory of underemployment will, in the following, be given a rather intensive analysis, in terms of an immanent criticism, for two reasons: (1) their importance. It seems worthwhile definitely to kill this whole approach, if it is misdirected; (2) the fact that they are representative of a much wider complex of concepts and theories in what the author calls the "modern approach" to the problems of underdevelopment, development and planning for development in underdeveloped countries.

1. THE CONCEPT IMPLIES A THEORY

The terms "underemployment" and its many synonyms—"hidden", "concealed", "invisible", "disguised", "potential", and "latent" unemployment—and "surplus" or "excess" labour, which together with open unemployment constitute the "labour reserve", are used with a considerable variation of meaning by different authors. Basically, however, these terms all represent an attempt to state the fact that the labour force actually engaged in a certain type of economic activity is idle during a part of the day, week, month and year, or, if working, is unproductive. To that extent, the time spent by the workers is useless and dispensable. The inference is drawn that a proportionally decreased labour force in a particular line of production should be able to produce as much as the actual labour force does, if the remaining workers worked full time. This inference establishes the theory that is implied in the concept of underemployment.

This waste of labour is "disguised" in the sense that the superfluous part of the labour force does not appear in the wage-labour market as workers seeking, but not finding, employment. The theory of underemployment mainly concerns only that portion of the labour force that is "self-employed", usually within a family enterprise.[2]

---

[1] This article is a draft of a methodological appendix in a book on South Asian development problems on which the author is working.

[2] To the "self-employed" in agriculture should be added those farm labourers who are permanently employed, sometimes called "attached workers." "It may be argued that a permanent farm servant, being a wage-earner, cannot be said to be underemployed. But there is an overwhelming evidence on overstaffing in all lines in the underdeveloped areas." (S. S. Gill, "Unemployment and Underemployment of Permanent Farm Workers," *Artha Vijñāna*, Vol. 2, No. 4, December, 1960, p. 249). "Here the employer has to employ a man even though half-a-man-worth work is all that he can give him." (N. V. Sovani, "Underemployment Micro and Macro and Development Planning," *The Indian Economic Journal*, Vol. II, April 1955, p. 303; cited by Gill, op. cit. Footnote 3, p. 249).

When, as occasionally, the term underemployment is used also for employed wage-labour, the waste of labour implied is assumed to be caused by low wages combined with "non-economic" motives on the part of the employers. Cf. Gustav Ranis and John C. H. Fei, "A Theory of Economic Development," *The American Economic Review*, Vol. LI, No. 4, September 1961, pp. 536 ff. Following W. Arthur Lewis' rightly famous article, "Economic Development with Unlimited Supplies of Labour," *The Manchester School of Economic and Social Studies*, XXII, No. 2, May 1954, the authors explain: "The persistence of this wage level is sustained by institutional or nonmarket forces since under competitive assumptions the real wage would fall to zero, at equality with . . .(marginal physical productivity). We shall call this the institutional wage."

Like most of the authors who have dealt with the problem of underemployment, we shall in the following be referring particularly to the labour force in the huge economic sector of small-scale, traditional farming. What we have to say, however, would be equally valid for self-employed workers engaged in petty trading and other non-agricultural, individual or family pursuits.[3]

It raises great difficulties to apply in South Asia the Western concept "unemployment" even to wage-labour; this statement we shall in the present context leave unproved. But for the "self-employed" the concept of unemployment is, of course, still less adequate. In this situation the concept underemployment is an attempt to find terms which permit a quantitative expression of actual idleness or unproductivity on the part of the "self-employed" workers. To the extent that they are idle or unproductive, those workers appear as a "surplus" labour force. In planning, underemployment is thus an adjunct concept to unemployment in the wage-labour market. Together, unemployment and underemployment make up the "labour reserve", available as a resource for planning. As Rosenstein-Rodan, who himself has made such outstanding contributions to the development of the theory of underemployment, observes: "Since the 1940's, it has been made one of the cornerstones of the theory of development of underdeveloped countries."[4]

---

[3] The following quotation from W. Arthur Lewis (*The Theory of Economic Growth*, George Allen and Unwin, Ltd., London, 1955, p. 326) gives a representative formulation of the total situation in a typical underdeveloped country, as seen by those authors who have elaborated the concept and theory of underemployment; it also notes the connection of this concept and theory with the earlier speculation in terms of over-population.

"This extreme state of over-population is, alas, not unknown. It can usually be recognized from the excessive numbers attached to certain sectors of the economy, especially domestic service, petty trade, casual employments and agriculture. Domestic service swells because such economies adjust themselves to the need for each person to provide as much employment as he can; social prestige requires each person to have servants if he can and the wealthier members of the community have to fill up their households with hordes of retainers who are little more than a drag on the purse. An extreme example of this is the island of Barbados, where sixteen per cent of the population is recorded by the census as engaged in domestic service. Petty trading shows similar extension; market places are crowded with stalls, in which each seller makes only a very few sales—selling and passing the time away in gossip being almost indistinguishable. Then there are the hordes of porters, jobbing gardeners, and others who get such casual employment as they can, averaging perhaps a day a week or less. In agriculture the phenomenon shows itself in the smallness of the farms; the plot cultivated by the average family is so small that it cannot fully occupy all the time of members of the family. Whether the excess of population shows itself mainly in agriculture, or mainly in domestic service, trading and casual jobs, depend on whether agriculture is operated with wage labour or by peasants. If it is operated with wage labour (e.g. Barbados) it will not employ more people than are needed for cultivation, and the surplus must then find a living outside agriculture. But if it is operated by peasants, the surplus lives on the family farms, and there may be very little surplus in other occupations. The normal tendency in over-populated countries is for the great landowners to let their lands to peasants for rent, rather than to hire agricultural labourers. They get more this way, since the wage they would have to pay to labourers is more than is left to the peasants when rents have been extracted from them."

[4] P. N. Rosenstein-Rodan, "Disguised Unemployment and Underemployment in Agriculture," FAO, *Monthly Bulletin of Agricultural Economics and Statistics*, Vol. VI, Nos. 7—8, July—August 1957, p. 1.

2. A PRECARIOUS ANALOGY

From the point of view of the history of economic doctrine, the concept and theory of underemployment is a recent development—steered in the particular direction which we have called the modern approach—of the old speculation in terms of over-population. The term and an essence of the theory as developed in recent years stems more directly from the Western concept "disguised unemployment" as introduced by Joan Robinson in the unemployment discussion during the Great Depression,[5] to denote workers who were pushed down into less productive and remunerative occupations as self-employed in order to have something to live on, or something more to live on than the dole, personal savings and family support. "It is natural to describe the adoption of inferior occupations by dismissed workers as *disguised unemployment*."[6]

The assumption underlying this concept as used in the Western milieu is, however, that this is a rather exceptional occurrence, relatively small in scope, and, more particularly, of a temporary nature. The "disguised unemployed" in the Western setting belong really to another, more regular and better paid, occupation where they will reappear as employed, or overtly unemployed workers, as soon as the labour market gives them a chance again to get jobs there; such a change of market conditions is normal and can be expected.[7] Except for demand for the products, which is temporarily deficient, all requisites for their employment in their regular occupation are present or will emerge when that demand has risen: on the one hand, enterprise, capital equipment and other necessities for production; on the other, skills and aptitudes, preparedness to step back, etc. All that is needed to eliminate disguised unemployment, as well as ordinary unemployment, is, therefore, an increase in the level of total effective demand for labour which, in its turn, is a function of aggregate demand for (capital and consumption) goods and services.

But when the concept underemployment—or its various synonyms, among them also frequently "disguised unemployment"—is used in context of the South Asian problem of labour utilisation, it refers to something entirely different. The vast and long-term underutilisation of human resources in which, permanently and structurally, more labourers are tied up in various lines of production than is necessary for producing the output. Those workers "belong" where they are, not in another more remunerative occupation. They never had another type of job from which they had been dismissed and to which they are expected to move back. Their vision of life cannot contain the idea of "return" to an occupation conceived by them as their regular and proper one. There is no potential employment for them once the

---

[5] Joan Robinson, "Disguised Unemployment", *Economic Journal*, XLVI, No. 182, June, 1936, p. 225 ff.

[6] Robinson, op. cit. p. 226. "He takes up some occupation—planting potatoes if he can get an allotment, selling match-boxes in the Strand, hanging round railway stations to carry bags to hotels." (Robinson, op. cit. 229–230).

[7] "If a revival of investment were to occur, dismissed workers would be called back from the hedgerows and the street-curbs into their normal occupations," (Robinson, op. cit. p. 227).

the spur of an increase in aggregate demand for goods and services becomes an effective demand for labour. The economies in South Asian countries are typically "low-elasticity economies," and an increased aggregate demand does not translate itself readily into an all-round increase in production and labour demand. And the under-employed themselves have not the skills, aptitudes and preparedness to move. Behind these fundamental differences lies the whole complex of less favourable conditions, which constitute underdevelopment, in regard to levels and modes of production and of living and to attitudes and institutions.

3. THE DEFINITION OF THE CONCEPT UNDEREMPLOYMENT

We can start out from the definition given by the experts who wrote the classic U.N. report, *Measures for the Economic Development of Under-Developed Countries.*[8] After having stressed that "rapid economic development is also the fundamental remedy for disguised unemployment" and that "the significance of the term 'disguised' is that it is applied only to persons who are not normally engaged in wage employment" but self-employed, the authors state: "The disguised unemployed are those persons who work on their own account and who are so numerous, relatively to the resources with which they work, that if a number of them were withdrawn for work in other sectors of the economy, the total output of the sector from which they were withdrawn would not be diminished even though no significant reorganization occurred in this sector, and no significant substitution of capital."[9] A similar line is taken in a great number of writings on the topic,[10] and is implied in practically the whole literature on underdeveloped countries, including earlier contributions by the present author. As already mentioned, the concept and theory of underemployment is an extension of the ideas of "overpopulation" and "population optimum" and, indeed, meant to give a clearer and a measurable expression of these ideas.[11]

---

[8] United Nations, Department of Economic Affairs, *Measures for the Economic Development of Under-Developed Countries,* (Report by a Group of Experts appointed by the Secretary-General of the United Nations), New York, May 1951.

[9] U.N. op, cit. p. 7. The experts continue: "The term is not applied to wage labour; presumably employers will not employ a labourer for wages unless his labour increases the total product.' (the implication being that marginal productivity is higher than zero; it would, however, need to be equal, or higher than, the wage; see Section 6 below on this question). "The use of the word 'unemployment' in this connexion is, therefore, somewhat misleading, since it is more often confined to wage-labourers whose status is recorded in unemployment statistics. We prefer to use, hereafter, the less precise but more familiar term 'underemployment' which is used in our terms of reference." The definition is repeated later in the report: "In many underdeveloped areas, the population on the land is so great that large numbers could be withdrawn from agriculture without any fall in agricultural output and with very little change of capital techniques." (U.N. op. cit. p. 41).

[10] "If we take the size of the labour force as given, underemployment may be described as a situation in which the withdrawal of a certain quantity of the factor labour to other uses, will not appreciably diminish the total output of the sector from which it is withdrawn." Alfredo Navarrete, Jr., and Ifigenia M. De Navarrete, "Underemployment in Underdeveloped Economies", *International Economic Papers,* No. 3, London, 1953, p. 235.

"These countries suffer from large-scale disguised unemployment in the sense that, even with unchanged techniques of agriculture, a large part of the population engaged in agriculture could be removed without reducing agricultural output. That is the definition of the concept of disguised unemployment as applied to the situation with which we are concerned. The same farm output could be got with a smaller labour force." (Ragnar Nurkse, *Problems of Capital Formation in Underdeveloped Countries,* Basil Blackwell, Oxford, 1953, p. 32).

One point which is often left unclarified in the definitions is whether the waste of labour is thought to consist of idleness on the part of the workers or of their working without producing anything, or both.[12] Sometimes the assumption is apparently that underemployment consists of a mixture of workers being idle and of their working unproductively when they are not idle, though the proportions are left unspecified. As underemployment commonly is assumed to imply a zero marginal productivity of labour, the alternative to idleness on the part of workers would anyhow be unproductive work. The relation between underemployment and productivity of labour will be discussed below in Section 6.

This whole trend of thinking in terms of underemployment observes the usual taboo in the modern approach to the problem of labour supply and labour input that only the quantitative dimension of labour is taken into account. Labour efficiency, i.e., labour intensity and skill at the present level of technology and capital intensity, is entirely outside the scope of study. Implicitly, labour efficiency is taken as given and constant, unchanged and unchangeable.

"Disguised unemployment may be defined as the part of the labour force which can be released from certain sectors of an economy *without changing the capital-intensity* and institutional framework, but by shifting some labourers to other productive work so that the remaining labourers may find full-time work (at least during that part of the year when season allows them to work), and further by improving the organization of work and division of labour, and also by introducing simple labour saving devices requiring little or no not addition to capital outlay." (Moti Lal Gupta, *Problems of Unemployment in India*, Nederlandsche Economische Hoogeschool te Rotterdam, 1955, p. 9).

"By structural under-employment we mean that the total agricultural production will not fall even if a particular number of agricultural workers are taken away from it and given employment elsewhere in the economy." (P. K. Mukherjee, *Economic Surveys in Underdeveloped Countries, A Study in Methodology*, Asia Publishing House, Bombay, 1959, p. 89).

Rosenstein-Rodan, however, speaking about "the basic concept, which has a clear and unequivocal meaning", defines it as 'that amount of population in agriculture which can be removed from it *without any change in the method of cultivation*, without leading to any reduction in output." (P. N. Rosenstein-Rodan, "Disguised Unemployment and Underemployment in Agriculture", *Monthly Bulletin of Agricultural Economics and Statistics*, FAO, Rome, Vol. VI, Nos. 7—8, July—August 1957, p. 1.

[11] Occasionally, the older terms are used as synonyms, cf., e.g., Stanislas Swianiewicz:

"Agricultural overpopulation means a situation in which there is, in a given region, a larger working population in agriculture than is necessary for the production of actual agricultural output by the prevailing methods. This means that a part of the labour may be withdrawn from agriculture, and this reorganization will not necessarily involve any change in the size and content of agricultural output. This definition implies the idea of a certain optimum of agricultural population at given conditions of place, time and agricultural technique." Stanislas Swianiewicz, "The Problem of Agricultural Overpopulation", *Economi dan Keuangan Indonesia*, Tahan Ke-11, No. 7, July 1958, pp. 330-331.

For a criticism of the theory of population optimum, see Gunnar Myrdal, *The Political Element in the Development of Economic Theory*, Routledge and Kegan Paul, London, 1953, pp. 38–39, and Gunnar Myrdal, *Population : A Problem for Democracy*, Harvard University Press, Cambridge, Mass., 1940, pp. 130 ff., particularly pp. 139 ff.

[12] Rosenstein-Rodan is clear; he assumes idleness, and so do most of those who, like him, have gone into actually "measuring" underemployment.

Again in conformity with the modern approach, it is also implied that underemployment, like unemployment, is entirely "involuntary", in the sense that the underemployed workers would want to work full-time and productively, had they only the opportunity to do so. This in its turn implies a market standardization of conditions for work and a rationality of attitudes to life and work such as in modern times have in Western countries become a social fact.[13] The very special sense in which "involuntariness" has to be understood will be discussed in Sections 4, 6, and 8.

This line of definition, exemplified above from the literature, seeks to establish the meaning of what is underemployment in the "static" sense: the number of workers is the only variable, while *ceteris paribus* is assumed in regard to everything else. But the comparison, could, of course, be made with a situation in which, by the application of certain policy measures, a country expects to change some of these "givens."[14] The usual explanation why the "static" concept has been adhered to —when an explanation is given and the "static' approach is not simply taken for granted—viz., that it is the only one that is clearcut and definite,[15] cannot be accepted.[16]

The more plausible explanation of why most authors, without any theoretical worries, have been led to apply the "static" concept is the force of theoretical conservatism, which is so strong in economic theory because of the metaphysical legacy of natural law and utilitarian philosophy and hedonistic psychology, of which it has never freed itself. In line with the traditions in economic theory, this approach offers an opportunity to reach a definition of underemployment that is not dependent upon a policy premise and thus appears to be purely "objective" and related only to the facts. It is supposed to give a definition of a "reserve" of unutilized labour that is available; the size of this reserve is assumed to be independent of whether anything in the policy field will be done about putting it into use, and independent of what these policy measures are. In Sections 8 and 9 below we shall take up for critical examination the logic of the "static" character of the concept and of the theory of underemployment. At this stage in our immanent criticism we shall not question it and assume no difficulty on this point.

---

[13] Rosenstein-Rodan with his usual greater clarity makes it an explicit assumption: "It is assumed that those who are in surplus are involuntarily unemployed." (Rosenstein-Rodan, *Monthly B.A.E.S.*, p. 3). This is always, however, implied by the other authors also.

[14] In defining underemployment it would, thus, of course, be possible to assume, instead, increased capital intensity in terms of farm implements, working capital of various types and permanent investments, improved techniques, and a changed institutional structure, particularly in regard to the size of the farm holdings. Many of the estimates made on more impressionistic lines often assume such changes. (See below, Section 4). The limit is represented by the unworldly statements, sometimes made, about the decrease in the labour force engaged in agriculture, to some fraction of its present size, that would be possible in a South Asian country, if agricultural production could be carried on in the same manner as in the United States or in Western Europe, i.e., in a completely different demographic, social and economic environment.

[15] Rosenstein-Rodan, who defines the purpose of this attempt as "one of clarifying the definition and describing the method by which surplus population in agriculture can be measured or estimated", (p. 1), states that the static concept is the basic concept, which has a clear and unequivocal meaning. The dynamic concepts, which he points to, "are full of pitfalls and are made less exact and certain in their meaning and results." (Rosenstein-Rodan, *Monthly B.A.E.S.*, ibid., p. 2).

[16] "Under very precise assumptions and in specific circumstances, a clear meaning can be given to the dynamic concept." (Rosenstein-Rodan, *Monthly B.A.E.S.*, ibid., p. 2). See below in Section 9.

4. MEASURING UNDEREMPLOYMENT

The definition of underemployment tries to delimit a magnitude which in principle—i.e., in so far as relevant data are ascertainable—can be measured. The important point to stress is that this definition assumes a normative element, viz. a standard of what full time work shall be assumed to be in terms of hours per day, days per week and weeks per year; the abstraction from the qualitative dimension of labour input implies, moreover, the further assumption that the actual labour efficiency, when working, can be taken for granted and assumed not to be changed or even changeable.

The assumption of a standard introduces an outsider's point of view into the world of facts. The normative element in the definition is implicit in the assumption, already noted, that idleness is "involuntary", i.e., that all the members of the labour force do want to have the opportunities to work full time. But in order to define and measure this idleness or inproductivity that is assumed to be "involuntary", it is, of course, necessary to make a definite further assumption about what full time work is. Idleness, whether "involuntary" or "voluntary," can be nothing other than the differences between this outside, normative standard, and the actual work performance, plus work performed that is unproductive, if this element of underemployment is also taken into consideration.

There are different ways of introducing this norm. The most direct way of doing it is to state the assumption in terms of a difference between the labourers available and of those "required" to do the work.[17]

Or the "surplus labour" is calculated as a difference between the labour forces on farms with different productivity, owing to more or less full labour utilization, with the more intensively cultivated ones—e.g., farms with double-cropping—taken as the norm. Sometimes the standard is derived from the productivity of labour on farms with the highest output per labourer, and the labour requirement translated to the less productive farms, which in turn gives a labour surplus when compared with the actually available number of workers. Often, particularly in general discussion, the methods of measuring underemployment have not been in the foreground, but only the idea that the same output could be produced by fewer workers even without any changes in capital input, technology, and institutional structure. The norm that workers should work "full time" and the assumption that their not doing it is involuntary are, however, basic to the concept of underemployment.

---

[17] Rosenstein-Rodan, *Monthly B.A.E.S.*, ibid., p. 3. This is not a simple matter, as differences have to be considered in regard to ownership relations, types of agriculture, seasons, climate generally, and health, age-structure, sex, etc., of the members of the labour force. See Rosenstein-Rodan, *Monthly B.A.E.S.*, ibid., pp. 2–5.

5. THE CAUSES OF UNDEREMPLOYMENT

As underemployment is assumed to be "involuntary", it has to be explained by the conditions under which the labourers are working which compel them, in spite of the rationalistic attitudes assumed, to waste their time in idleness or unproductive work.

In line with the classical tradition in economic theory, this explanation is regularly stated in terms of insufficiency of complementary resources and a low level of the arts, with stress usually on the first factor. The U.N. experts quoted above and other authors as well write this explanation into the definition itself. In regard to the first factor the U.N. experts further explain:

> "Under-employment is due to a deficiency of the resources which are necessary to employ productively the available supply of labour. It is usually associated with family employment where, in agriculture or industry, the unit of production and of the supply of labour is the family; it exists because the resources of the family are too small to keep all working members of the family fully employed throughout the year and because there exist no opportunities for directing a part of the supply away into other occupations at appropriate times."[18]

As the opportunities for employment outside agriculture presumably also depend on available resources there, lack of resources in the national economy as a whole simultaneously both decreases the work opportunities in agriculture and the demand for labour outside agriculture, in both ways keeping up the volume of underemployment in agriculture.[19] Similar relationships exist naturally in regard to the second factor, the level of the arts. Particularly as resources are generally translated into capital, the theory of underemployment is thus fit into the usual economic development models where an increase in output is made a function of capital input with reservation for the support rendered by progress in the arts.

There are, however, two snags in this general line of explanation. First of all it might be the case—as the U.N. experts assume—that lack of complementary resources (and presumably also primitive techniques) decreases the productivity of a greater labour input in agriculture, or even the possibility of an additional labour input; but the effect of raising levels of capital intensity and technology might be to release labour at an optimal level of production. This is, in fact, usually assumed to be the case, particularly when the changes in methods of production in agriculture amount to greater mechanization.[20] But the theory of underemployment is stated in static terms of unchanged levels of capital equipment and technology. The point has importance only when it is used in a consideration of policy, which for realism must be in dynamic terms.[21]

---

[18] U.N., op. cit. p. 8. Cf. Government of India, Planning Commission, *The First Five Year Plan, A Draft Outline*, New Delhi, 1951, p. 19: "Corresponding to idle labour, there are no adequate supplies of other cooperating factors of production, such as land and capital equipment."

[19] Occasionally the stress is put only on the latter causal relation: "The primary cause of underemployment in the non-industrial sector is the lack of demand for labour emanating from the industrial sector owing to the absence of complementary factors to enable industrial expansion." (N. Islam, "External Economies and the Doctrine of Balanced Growth," *Indian Economic Journal*, July, 1957, p. 51).

[20] In the short run, the degree of mechanization that is at all practicable may on the whole rather be an opportunity for bigger productive labour input in agriculture. If this were so, there would be no hitch on this point.

The second hitch is that, already for logical reasons, some change in the way factors of production are combined must be assumed.[22] The general line is — explicitly or, mostly, implicitly — to make a distinction between, on the one hand, technological advance and capital investment—which are understood to go together — and, on the other hand, the organization of production and work. The former category of changes are excluded by definition, while a certain change of the latter type must be obtained, though without major changes of the institutional framework.[23]

---

[21] Nurkse gives a particular reason why he believes that the static assumptions make sense, at least in the first approach to the problem: "Changes in technical methods are excluded from the definition of disguised unemployment. Improvements in methods are extremely important. Experts seem to be agreed, however, that it is rather hopeless to try to introduce better farming methods unless the excess population is drained off first. There is little chance of any substantial advance in agricultural technique until some of the factors of production now engaged in that activity have been removed. This may sound paradoxical, but there is some basis for this view." (Nurkse, *Problems of Capital Formation*, p. 34).

[22] This has been seen by Rosenstein-Rodan: "If one adheres rigidly to the assumption of *ceteris paribus*, under which nothing whatever can be changed in the method of cultivation, even a mere one-month peak load of work can substantially lower the amount of disguised unemployment. It is, however, convenient and also reasonable to deviate slightly from this assumption of no change, provided that the extent and the nature of the minimum change deemed compatible with static assumptions is described in detail. This is the same method as that which is applied in the theory of supply, where a minimum amount of imperfect competition is always assumed." (Rosenstein-Rodan, *Monthly B.A.E.S.*, ibid., p. 4). With this motivation, he allows in his Italian enquiry that workers working for less than 51 days are assumed to be "removable", which naturally, in its turn, assumes some change of organization; in his study the change consists in the employment of additional labour from the outside labour market.

[23] "We exclude technological advance, more equipment, mechanization, better seeds, improvements in drainage, irrigation, and other such conditions. We exclude these things here only in order to isolate the possibilities that stem from the presence of large-scale disguised unemployment. There is, of course, no question of excluding them from any development programme in practice.

One thing, however, one need not and probably cannot exclude and that is better organization. If the surplus labour is withdrawn from the land the remaining people will not go on working in quite the same way. We may have to allow for changes in the manner and organization of work, including possibly a consolidation of scattered strips and plots of land."(Nurkse,*Problems of Capital Formation*, p.33).

Everett E. Hagen follows the same line: "By a strict definition, the term underemployment refers (in agriculture) to a condition in which, if some workers were removed from agriculture, the same total output could be produced by the remaining workers without any change in methods. I think it probable that this situation, strictly defined in this way, is fairly rare. Methods are in effect that make some use of all available labour. But if we modify the definition to what might be termed a practical one, namely that if some agricultural workers were withdrawn, the same total output could be produced by the remaining workers after a relatively simple adaptation of methods, then I think there can be little question that there is a large amount of agricultural underemployment in China, Japan, India, Ceylon, Malaya, Indonesia, Egypt, countries of southern and eastern Europe and a number of other countries." (Everett E. Hagen, "Economics and Economic Development," in *Economics and the Policy Maker, Brookings Lectures*, 1958–1959, The Brookings Institution, Washington, D.C., 1959, pp. 193–194).

Also, V. V. Bhatt, "Underemployment in Underdeveloped Economies: Theoretical Considerations," *Indian Economic Journal*, July 1957, p. 44: "*With appropriate reorganization*, the output of agriculture will not be significantly affected and in some cases may even increase by removing a certain quantity of labour from it." (my italics).

See also Gupta's definition of underemployment, quoted above in a footnote to Section 2, which spells out the assumption of such changes: "...by introducing simple labour saving devices requiring little or no net addition to capital outlay." (Gupta, *Problems of Unemployment*, p. 9).

All writers who have gone deeper into the theory of underemployment have, of course, given consideration to the fact that some and in most cases a considerable part of underemployment is seasonal and caused by the cyclical changes during a year of the need for labour input. In addition, Rosenstein-Rodan also observes that when the units of production are small, it becomes of importance that part of the underemployment is "fractional", defined as "labour hours not used throughout the whole year which do not add up to an entire labour unit."[24]

In both cases there is true underemployment, though its utilization is limited spatially, as the new work opportunities which could absorb it cannot be far away from the farms where the underemployed now live and work part of the time.[25]

The modern theory of underemployment is, as we noted, a development of the old speculations about overpopulation. So far as the historical explanation of underemployment and the prospect for the future are concerned, most authors refer to the population increase, though this factor, naturally, does not enter technically in the theory, which is static and by definition refers to a point of time. The assumption that, when there is underemployment, marginal net productivity of labour is zero (see below Section 6) implies, however, that an increase in the labour force, following a population increase, will not raise output but only increase underemployment.

6. THE IMPLICATIONS IN REGARD TO MARGINAL PRODUCTIVITY

Practically all authors make the point, usually in direct connection with the definition of underemployment, that the definition implies that marginal productivity of labour is zero. This is apparently assumed to be self-evident and, therefore, never argued.[26]

---

[24] Rosenstein-Rodan, *Monthly B.A.E.S.*, ibid., p. 4. In Rosenstein-Rodan's calculation for Southern Italy about one half of total underemployment was considered "non-removable" on account of its character of being seasonal or fractional.

[25] "Even where disguised unemployment is mainly a seasonal matter, the question of making productive use of it still arises and still has important implications in regard to capital formation." (Nurkse, *Problems of Capital Formation*, pp. 35–36).

"Persons in fractional disguised unemployment cannot be moved out of agriculture, though they could be provided with more part-time work in handicrafts, community development, etc." (Rosenstein-Rodan, *Monthly B.A.E.S.*, ibid., p. 4).

[26] Rosenstein-Rodan, *Monthly B.A.E.S.*, ibid., p.1 : "The marginal productivity of labour, *in other words*, is zero." (my italics)

Navarrete and de Navarrete, *Int'l Econ. Papers*, 3, p. 235; "This is as much as to say that the marginal productivity of these units of the factor labour in their original employment is zero, or very close to zero." It is not explained why marginal productivity is not unreservedly zero, if that is implied.

Nurkse is occasionally a little more careful, and thus unclear: "And in this broad dynamic sense the marginal productivity of labour can perhaps be said to be negative." (Nurkse, *Problems of Capital Formation*, p. 34). On the preceding page he states, however, referring to an overpopulated country, where the labour-force on the land could be dispensed with without making any difference to the volume of output: "In technical terms, the marginal productivity of labour, over a wide range, is zero."

Some authors take the more extravagant position that the marginal productivity of labour might be negative, which is supposed to imply that—under the ordinary static assumptions in regard to capital, technology and the institutional framework—output would increase if the underemployed were removed[27] and, in fact, even without removal if they decreased their labour input. Others imply in regard to the marginal productivity of labour that it may be positive but is very low.[28] We may note as a start that it is somewhat disturbing to have a corollary to the definition, which, though it is so self-evident that it is never argued, nevertheless turns out to hover uncertainly between the limits of "very low" and "negative".

The central position in the theory of underemployment of the thesis that marginal productivity is zero—or very low or negative—is indicated by the fact that those few authors who lately have been critical of the theory have directed their attack on this thesis, while apparently accepting its logical relation to the concept underemployment.[29]

---

[27] "As the normal agricultural output can be obtained with a smaller labour force than is actually employed, the marginal productivity of the latent surplus of labour is zero, most probably, even negative. This implies that by removing some workers, agricultural output can actually be increased." (Gupta, *Problems of Unemployment*, p. 24).

"When a country is over-populated, on this definition, [so as to cause disguised unemployment in agriculture] the marginal productivity of labour in agriculture is zero. It may even be negative. The Law of Diminishing Returns asserts that it is possible to have too many people on too little land, and this is not an uncommon phenomenon." (Lewis, *Aspects of Industrialisation*, National Bank of Egypt, Fiftieth Anniversary Commemoration Lectures, Cairo, 1953, p. 8).

Nurkse makes the point, however: "Some observers suggest that it may even be negative, which would imply that, by removing some people, farm output could actually be increased. The reason for this might be that under existing conditions people actually get into each other's way, so that if some go away those who remain are able to work more effectively. But this seems to me a doubtful and, in any case, unnecessary assumption, and I am not going to use it." (Nurkse, *Problems of Capital Formation*, p. 33).

[28] Occasionally, under the influence of reminiscences from the Malthusian and classical theory of population and wages, marginal productivity is related to the subsistence level.

". . . even at the existing level of technique there is redundant agricultural labour in the sense that the marginal product of labour is less than the minimum subsistence requirements of the workers. Since farming in most of these countries is largely on a family basis, the redundance of agricultural labour finds expression not so much in the existence of mass unemployment as in the phenomenon of chronic underemployment affecting the whole population." (United Nations, ECAFE, *Economic Survey of Asia and the Far East* 1950, New York, 1951, p. 66). Presumably the meaning is that people in family enterprises are rewarded according to average product, which, of course, should be higher than marginal product, which then could fall under the minimum subsistence level. The lack of logical clarity in this reasoning is obvious.

[29] "The term 'disguised unemployment' is commonly used to designate a situation in which the removal from a working combination of factors of some units of labour, nothing else of consequence or worth mentioning being changed, will leave the aggregate product of the working combination undiminished, and may even increase it. To say that there is 'disguised unemployment' is therefore equivalent to saying that in that working combination the marginal productivity is zero or almost zero and may even be a negative quantity. (Jacob Viner, *Contribuicoes a Analise do Desenvalvimento Economica*, p. 345).

By demonstrating that marginal productivity is not, or cannot be, zero, they believe that they have disposed of the concept of underemployment.

The astonishing thing is that in this large literature on underemployment, and also in the more recent controversy just referred to, little attempt has been made to clarify what, more particularly, is meant by marginal productivity, and why zero marginal productivity—or something near it—is a condition for underemployment.

Most authors seem to have in mind a situation where, under otherwise unchanged conditions, an addition to the labour force would result in a corresponding increase in idleness[30] and, consequently, in a sharing of the workload with no addition to the actual labour input and no addition to output. In the other direction the labour force could then be decreased without lowering output, until all idleness would have been liquidated. This could, perhaps, be described as a *zero marginal productivity of an addition to the labour force.* But, firstly, it is not understandable why this situation is described with reference to a margin, as it would reign not only at the actual margin but far back on the scale of the number of workers available—until a point when the workers were so few that there would be no idleness anymore.

Secondly, productivity must logically be related not to the number of workers, but to their labour input. Aggregate labour input, however, is by assumption held constant—as a corollary to the assumption of constant output under unchanged conditions in terms of capital input, production techniques and institutional conditions. Under this assumption the workers' *average productivity* is naturally a function—and a very simple one—of the size of the labour force. But what sense does it make to state that their *marginal productivity* is zero, when this only expresses the assumption that an increase in the labour force does not imply an increase in the labour input?

The concept marginal productivity of labour must, of course, logically refer to the labour input, which in the ordinary way must be reckoned not in terms of the size of the labour force, whether it is working or not, but in hours, days and weeks of work actually performed. If that is clear, it cannot be taken for granted that the fact of underemployment in terms of actual idleness on the part of the labour force implies a zero marginal productivity of labour. On the contrary there could,

---

"A large literature has appeared in recent years based on the belief that countries, however poor the collection of resources at their disposal, can do wonders by recombining the resources already at hand. The mainspring of this view has been the notion that in agriculture one finds many workers whose marginal productivity is not only far below that in other sectors but that it is zero, and, of course, the agricultural sector bulks large in nearly all poor countries." (Theodore W. Schultz, "The Role of Government in Promoting Economic Growth," in Leonard D. White (ed.), *The State of the Social Sciences*, University of Chicago Press, 1956, p. 375).

[30] "In agriculture the phenomenon shows itself in the smallness of the farms; the plot cultivated by the average family is so small that it cannot fully occupy all the time of members of the family." (W. A. Lewis, *Theory of Econ. Growth*, p. 326).

of course, and normally should, be underemployment according to the definition given — and provided that this definition and the theory implied were logically consistent and realistic in other respects — at a marginal productivity of labour higher than zero. How high, would depend upon the level of marginal productivity, at which the workers would not find it worth while to put in an extra unit of labour.

The whole discussion of marginal productivity in connection with workers' willingness to exert themselves in South Asian agriculture is airy and unrealistic; but this is so because of the assumption of rationality in attitudes, which has it most flagrant expression in the assumption that underemployment is "involuntary." If in our immanent criticism we accept this commonly made assumption, it is certainly not warranted to assume that underemployment implies zero marginal productivity of labour. Why should workers continue to work though they are not producing anything? And if they do, why don't they go further, so that there is no idleness at all? More particularly, why don't they stop working at a level above zero marginal productivity? Is it assumed that work does not represent any disutility?

As was mentioned in Section 3, one sometimes finds allusions to another conception of underemployment *viz.*, that labour is not productive.[81] Again, the question must be raised: how can the authors, who use the modern approach and, more particularly, assume that underemployment is involuntary in a very specific and rationalistic sense, also assume that workers behave as if work performance had no element of "disutility"? Moreover, how can a situation in which workers are working unproductively be characterized in terms of a zero marginal productivity, when this situation is meant to apply far inside the margin? It should be added that this construction does not make more sense in the mixed case, where underemployment is supposed to consist of both idleness and useless work.

---

[81] M. A. Teland, "Technique of Measuring Rural Unemployment," *Indian Journal of Agricultural Economics*, IX, No. 1, March, 1954, p. 150:

"This—disguised underemployment — includes the category of persons whose employment for a particular period is not really necessary ... . The farmer with a small piece of land is found to spend almost as much time on his land as a farmer with a larger piece of land because of lack of other employment opportunities he feels like doing some work during the day on his farm . . . All this work is not necessary, all the same it does occupy a great deal of a person's time. In any enquiry such persons will never call themselves unemployed."

Rosenstein-Rodan makes it clear that this is not his conception, but that underemployment with him has its reality in actual idleness. For most authors it is less easy to lay bare what their understanding of the situation is. The clearcut formulation of the second conception is given by some of the critics of the underemployment theory, as when Viner points out that "The 'unemployment' in the term 'disguised unemployment' — may be only metaphorical, since there may be hard work even at the margin, where 'unemployed' must mean 'unproductively employed'." (Viner, *Contribuicoes*, p. 345). Cf. p. 351: "But for Lewis, the 'disguised unemployed' of agriculture may be working as hard as anyone else, and many in fact need more calories for farm-work than they would need for factory-work. They are 'idle' in the sense only that their work is unproductive. (p. 351).

The conclusions we reach are the following. The statement regularly made, that underemployment as defined implies zero marginal productivity of labour, makes no sense. It does not stand, even when the assumption of rationality, on which it is based, is granted. The question then is, why this statement is always made, and why, furthermore, there are no attempts to clarify its meaning and to produce evidence why it is, or can be, a valid generalisation.

The only explanation for this peculiar situation is that this is theorizing within the blinkers of an ideology, in this case what we have called the modern approach, borrowed from the West and corresponding to Western conditions.

7. THE ACTUAL LEVEL OF MARGINAL PRODUCTIVITY IN UNDERDEVELOPED COUNTRIES

Under these circumstances, it is natural that the critics of the underemployment theory pick out the assertion about marginal productivity being zero for their criticism. We should note though, that when they succeed in throwing doubt on the assumption of zero marginal productivity of labour they have not disposed of that theory. For, as we have shown, there could be underemployment at a higher than zero marginal productivity of labour.

So long as the idea of zero or near-zero marginal productivity of labour is widespread, and plays a large rolė even outside the formal theory of underemployment, the following observations are in place.

Viner makes the point that zero marginal productivity would assume fixed technical coefficients "in a valid economic sense" and fixed crop patterns or uniform, fixed technical coefficients for all products. On logical grounds he denies that this is easily conceivable, and, on common sense grounds, that it is likely ever to be the case in any country and any occupation.[32]

The middle position—where there is a mixture of idleness and low productivity—is characterized by another critic of the underemployment theory, Schultz: "The idea that disguised unemployment is characteristic of poor countries is very widely held. It has become one of those self-evident propositions. Presumably anyone who has ever been about in poor countries is prepared to testify from his observations that it is true. Many people, so the proposition runs, are really working only part-time, doing little productive work, some work only seasonally, and others while they are employed most of the year are not kept busy, or, if they are, they are busy doing very little that contributes to production." (Theodore W. Schultz, *The Economic Test in Latin America*, New York State School of Industrial and Labour Relations, Bulletin 35, Cornell University, Ithaca, New York, 1956, pp. 13, 14).

[32] "As far as agriculture is concerned, I find it impossible to conceive a farm of any kind on which, other factors of production being held constant in quantity, and even in form as well, it would not be possible, by known methods to obtain some addition to the crop by using additional labour in more careful selection and planting of the seeds, more intensive weeding, cultivation, thinning, and mulching, more painstaking harvesting, gleaning and cleaning of the crop. Even supposing that there were such a farm on which every product had technically and economically fixed ingredients, labour would still have positive marginal productivity unless there were not only fixed technical coefficients of production for all the economically relevant potential products of the farm, but the proportions between the technical coefficients were uniform for all of these products. For, if these proportions are different as between different products, then it will always be possible by appropriate change in the product mix, in the direction of more production of those products whose labour technical coefficients are relatively high, to absorb productively any increment of labour." (Viner, *Contribuicoes*, p. 347).

He also points to the fact that where the family labour supply is not adequate or is unable to work, labour is employed for wages, which is not compatible with a zero marginal productivity of labour.[33]

These critical points[34] seem to be incontrovertible, so far as they go. In regard to the first argument, concerning technical coefficients and crop pattern, it must again be recalled that the whole discussion of underemployment is a continuation of the older discussion of overpopulation caused by population growth. What actually happened during this development towards an ever higher man/land ratio illustrates in fact how very wide the possibilities are of varying the labour intensity of cultivation — and thus the total labour input per hectare — even within a given cropping pattern and without appreciable change in techniques. One consequence of population growth has very generally been a movement to more intensive utilization of the land resources. The same possibilities are demonstrated by the huge differences in the labour intensity in paddy production in the South Asian region. Likewise, the methods of irrigation and the utilization of irrigation—which does not signify a new development of agricultural technology—can vary widely in response to changes in needs and in availability of labour. And, of course, the cropping patterns have been varied and can be varied, especially by adding to, or substituting for, the usual subsistence crops other and more labour-intensive products, such as sugar, jute, potatoes and other horticultural products. Far from being determined rigidly by natural conditions, or by the institutional framework, the techniques and the capital implements, the amount of labour applied to a unit of land is highly variable in response to changes in the intensity of population pressure.

To this should be added the fact that in South Asia, generally speaking, the yeilds per hectare are mostly very low outside the plantation sector. There is no doubt that, even under the assumptions made in the theory of underemployment about unchanged institutional framework, capital equipment and techniques, the yield can almost always be raised by an increased labour input. The marginal productivity is higher than zero.

---

[33] "Unless one assumes non-economic motivation on the part of employers, there is difficulty in conceiving why they should hire at any wage-rate additional units of labour beyond the point at which they know the labour will add less in value to the product than the wage-cost, to say nothing of the case where the labour will add nothing to and may even subtract from the product." (Viner, *Contribuicoes*, p. 344). Viner's assumption of rational economic motivation is justified in this particular argument, as the reasoning he is criticising is carried out in terms of such motivation.

[34] Schultz's criticism is less closely argued. He merely asserts that "The marginal productivity of labor in agriculture in poor countries is very low because of the poor collection of resources, but it is not zero. Moreover, in such a country it is very low for labor generally." He adds the information, without developing or proving it, that "and in situations, and there are many, where for many years, often for decades, agriculture has approximated a kind of stationary state, one is likely to find the average and the marginal values of labor more nearly the same than would be the case in a rapidly developing economy like that of the United States." "I know of no evidence for any poor country anywhere that would even suggest that a transfer of some small fraction, say, 5 per cent, of the existing labor force out of agriculture, with other things equal, could be made without reducing its production." "A poor country which has been virtually stationary for a long period is likely to reveal any appreciable malallocation of factors, say, as between agriculture and the rest of the economy or within agriculture, whereas a country undergoing rapid economic growth, more likely than not, will have in it areas that have been bypassed and others that have become depressed as a consequence of its economic development." (Schultz, *Role of Government*, p. 375–376).

But naturally both aggregate and marginal productivity are low in poor countries and particularly in traditional self-sufficiency farming. As Viner puts it: "there is little in all the phenomena designated as "disguised unemployment', as 'hidden employment', or as 'underemployment' which in so far as they constitute genuine social problems would not be adequately taken into account by competent, informed, and comprehensive analysis of the phenomenon of low productivity of employed labour, its causes, its true extent and its possible remedies."[35]

8. A "STATIC" THEORY ABOUT THE RESULTS OF A BIG CHANGE

The concept and theory of underemployment are developed in static terms by assuming unchanged conditions in regard to capital input, techniques and institutional framework. The assumption of minor elements of change in organization of work,[36] necessary to carry out the analysis, does not by itself seriously infringe upon its static character. It is carried out, however, by a comparison between two situations, both fulfilling the foregoing assumption of unchanged conditions, but separated by a very big change, *viz.*, the decrease of the labour force. Underemployment is measured by the relative size of that part of the labour force that can be "withdrawn", "released" or "removed", without causing a decrease in production. This change is thought of in purely abstract terms.[37] It is also timeless, as no assumptions are made in regard to a period of time during which the change should take place. The question is then, however, what sense, in pure logic, a "static" theory about the results of a very big change can have.

---

[35] J. Viner, *Contribuicoes*, p. 354.

[36] See above in Section 5.

[37] "A number of people are working on farms or small peasant plots, contributing virtually nothing to output, but subsisting on a share of their family's real income. *There is no possibility of personal identification here*, as there is in open industrial unemployment. In industrial countries unemployment is a glaring waste of resources, visible to all, and has perhaps for this reason attracted more attention. In an overpopulated peasant economy, *we cannot point to any person* and say he is unemployed in disguise. The people may all be occupied, and no one may consider himself idle." (my italics). (Nurkse, *Problems of Capital Formation*, p. 33).

". . . the definition of chronic underemployment in macro terms considers not individual workers but abstract labour units and the extent of underemployment indicated by it is that of so many man-hours." (N. V. Sovani, "Underemployment Micro and Macro and Development Planning," *Indian Economic Journal*, Vol. II, April, 1955, p. 304). Sovani, who is sceptical about the concept, notes that "such a conversion seems to present no difficulty to many". (Sovani, op. cit. p. 308).

The change is of a truly revolutionary magnitude, ordinarily assumed to affect from twenty to fifty per cent of the labour force. If it should happen, it could not be abstract any longer, but must mean that a great many people move. It can only take place in a period of time, shorter or longer, and the length of the time allowed for the change cannot be without influence on everything, including output. It is not legitimate logically to assume that conditions, which *must* change in the process, will remain unchanged.

The change does not happen by itself but must be thought of as the effect of government policy. Can the government ever be assumed to undertake such a policy while observing *status quo* in regard to capital input, techniques and institutional framo-work? Indeed, what meaning can be given to a theory assuming this? It should be added that the fact that traditional economic theory, as it is often expounded, comprises many other concepts and theorems of the same doubtful character, sometimes called "comparative statics", cannot be invoked to support this particular one. They may all be meaningless.

Let us assume, however, for the sake of the argument, that this static approach with an implied, abstract and timeless, immense change restricted only to the "removal" of workers, can be given a meaning. The next question is: how can it be assumed, as is assumed, that the reduced labour force will exert itself and so keep up the production at the same level as it had before the removal of the underemployed? Only so could production remain unchanged under the assumed static condition in regard to capital input, techniques and institutional framework. This, in turn, assumes that the remaining workers would suddenly begin to increase their labour input so much that there would no longer be any idleness. That they behave in this way is apparently the more specific meaning of the assumption — sometimes explicit but mostly only implied — that underemployment is "involuntary." As we pointed out in Section 4, the concept full employment and, consequently, idleness, in this analysis is injected as an outside norm which is revealed whenever an attempt is made actually to measure underemployment.

Is this assumption about the "involuntariness" of idleness — defined in this very specific way — realistic? Can it be realistic in any society? To the extent that idleness is decreased, the workers remaining in agriculture would enjoy higher real incomes after the removal of a part of the labour force. In the extreme case assumed by this theory, the rise in incomes would be very substantial, amounting to 50 per cent if, for instance, underemployment is supposed to be a third of the labour force. Would this rise in incomes have no effect on their ability to work, and their willingness to work?

In the first respect — ability to work — it must be expected that in poor countries, where nutrition and other levels of living are sub-optimal from the productivity point of view, the change would raise the quality of labour, when work is performed, and thus increase the productivity of the labour unit. But the whole approach to the employment problem, typified by the concept and theory of underemployment, is characterized by the neglect of labour efficiency.[38] If this causal relation between levels of income and labour efficiency were taken into account, and if the remaining workers began to work full time, which the theory assumes, the total produce would increase, which is against one of the assumptions. If then, instead, unchanged production were assumed, a net of idleness would be left, corresponding to the rise in labour efficiency. From the point of view of the theory of underemployment, this would, in fact, mean the possibility of a still higher "removal" than corresponds to the actual idleness. Underemployment would have been underestimated. The theory is open, because its two assumptions (1) liquidation of idleness —supposed to be "involuntary"—and (2) unchanged production, are incompatible.[39]

In the second respect, i.e., in regard to willingness to work, any element of survival-mindedness on the part of the workers, of which there is much in the South Asian region, would clearly motivate them not to begin working full time, when they could get the same income by sticking to work-habits which are customary. Aggregate production would then fall, though of course not production per head. If, in addition, there were decreasing returns, the remaining workers would be in the position to preserve their customary income levels and at the same time work even less than before the removal of a part of the labour force. From the point of view of the underemployment theory, this would imply that after the removal of a part of the labour force corresponding to the actual idleness, there would continue to be underemployment and, in fact, a relatively large underemployment.

This is an extreme case. But, more generally, it is an entirely unrealistic conception of the conditions for, and the attitudes to, life and work in the region to take it for granted that a removal of part of the labour force, corresponding to the amount of actual idleness—defined by the utilisation of an arbitrary norm representing an outsider's view—would motivate the remaining workers to work full time in order to keep up aggregate production at unchanged levels, and also that income levels have no effect upon labour efficiency.

The upshot of all this is that the static concept and theory of underemployment is a useless construct. In spite of the many thousand pages devoted to their exposition by so many distinguished authors, and the calculations and estimates of actual underemployment carried out with sometimes considerable precision, this concept and this theory suffer from lack of clarity at crucial points. When the points are clarified, as we have tried to do in this immanent criticism, they are no longer tenable.

---

[38] See above Section 3.

[39] Leibenstein has touched on this problem, though only from the point of view of why wages can be positive when marginal productivity is zero; see Harvey Leibenstein, *Economic Backwardness and Economic Growth*, John Wiley & Sons, Inc., New York, 1957, pp. 58ff. *at passim*.

When underemployment is measured, the only accomplishment is to have ascertained the amount of actual idleness. In order to do this, a norm has to be introduced, defining what is full employment. This norm represents the outsider's view of how much people ought to work, not the view of the people concerned. In spite of the arbitrariness of this norm applied in the measurement of idleness, this exercise, by itself, is clearly of considerable interest.

But the further manipulations — assuming static conditions in regard to capital input, techniques and institutional framework; assuming the attainment of full employment, when a part of the labour force, corresponding to the amount of idleness ascertained, is removed; assuming that this implies unchanged production; and putting it all in the term of a static comparison implying the vast change of removal to be instantaneous and not disturbing any of the above assumptions — does not introduce any new elements of knowledge but a number of abstract notions which, when spelled out, show themselves to be totally inadequate to reality and, therefore, invalid. These manipulations, which constitute the preliminaries to the definition of underemployment as that part of the labour force which under these assumptions can be removed and, which, together with open unemployment, represents the labour reserve, constitute the theory of underemployment. The real function of that theory is to make it possible to define underemployment, with that implied policy relevance, without resort to a policy premise, and, at the same time, to give credulence to the preconceived, but seldom explicitly stated, ideas — which represent what we have called the modern approach — that all idleness is "involuntary" in a specific and extreme sense and that labour can be dealt with in purely quantitative terms, disregarding the qualitative dimension.

Both the endeavour to escape stating the value premises for a study and the particular preconceived ideas we have discussed are widely prevalent as well in contexts other than that of underemployment, and this has motivated us to use the occasion for an illustrative demonstration.

9. AN ESCAPIST APPROACH

Much more intensive empirical research concerning the actual utilization of the labour force in the various countries of the region and the conditions determining their work is needed to make our knowledge fuller and more correct. Meanwhile, however, on the theoretical level we can at least clarify what we are actually talking about, and what questions we can raise. This is necessary in order to render the empirical research more adequate and pertinent. There is no, and can be no, labour reserve *in abstracto*, independent of the time allowed, the foreseen changes and particularly government policies.[40]

---

[40] ". . . any estimate of the removable surplus is really inherent in the assumptions one makes regarding the existing situation and the change in it. A removable population surplus is not something which absolutely and independently exists but is what we create even theoretically by assuming certain specified changes. Any estimate of it in any particular context are all equally valid under their own assumptions." (N. V. Sovani, "Underemployment, Movable Surplus and the Saving Fund," *Artha Vijñana*, Vol. 1, No. 1, March, 1959, p. 25).

To begin with, a change in labour input must be thought of in dynamic terms and must be related to a definite period of time. Labour input, when it is self-employed, as well as the supply and demand of wage-labour, has a qualitative aspect as well, labour efficiency, that must be considered in the dynamic process. The utilization of the labour force will in the period under consideration be dependent on changes that occur, which are all interdependent in their causation and of which some are changes of the workers' attitudes to life and work. Among the strategic changes, from a planning perspective, are government policies operating by means of education and compulsion, subsidies, taxation, investments, land reform, etc. Output will, of course, not be unchanged, and the changed output belongs among the changes in the social system the effect of which will have to be taken into account. Looked at *ex ante*, an increased labour utilization expected to be reached in a period of time by the application of government policies adapted to the specific —and changing—circumstances in a particular country may be seen as a "labour reserve" or "underemployment" to be cured. But these terms are by their present use so loaded with false association that we had better think in terms of utilization of the labour force—though, of course, the actual and expected partial idleness of workers is an important factor among many others.

The main point is that statements about the possibility of a better utilization of the labour force, if made without a policy assumption, are senseless—or, if we want to be charitable, we can say that a policy assumption is made though it is not spelled out. When, for instance, Schultz asserts that he knows "of no evidence for any poor country anywhere that would even suggest that a transfer of some small fraction, say, 5 per cent, of the existing labour force out of agriculture, with other things equal, could be made without reducing its production", he might be right in a sense. But then he must have assumed either such a short-term perspective that policies of the above-mentioned nature have little chance of changing conditions very much, or that, even in the longer view, policies will not be of much importance: because the government follows a *laissez-faire* line in regard to economic development in general, because its development policies, for whatever reasons, are not specifically directed towards increasing and improving labour utilization or because, in any case, the social setting—including workers' attitudes—are such that it will prevent its policies from having much effect on average labour input and labour efficiency. Such policy assumptions are perfectly feasible and may be realistic in many countries; but they need to be spelled out and their realism in the particular case needs to be established.

The opposite, and more common, view that the utilization of the labour force can be substantially increased and that, thus, "underemployment" is very big, would instead assume a vigorous development policy, which, overcoming vested interests, administrative bottlenecks, adverse ingrained institutional and attitudinal conditions amongst the population and other factors of social inertia, is directed towards accomplishing the better utilization of the labour force, and would, in addition, assume sufficient time to reach effect. Such assumptions, of course, can be made. But again, they need to be stated as a basis for the calculations. As these assumptions are anything but simple and self-evident, they need to be spelled out in specific terms. And to render the calculations at all realistic, it needs also to be established that the policy assumption are reasonable in the particular country under discussion.

By logical necessity, the main thing is that the concept "underemployment" is dependent upon assumptions about policy. As a rule neither party in the controversy referred to above has spelled out their assumptions in this regard. The concept is left floating and the arguments don't meet. It is quite possible that both parties may be "right": that in the short run, and also in the long run, in the absence of a vigorous development policy directed toward "creating employment", stimulating labour input and supply and raising labour efficiency, there might not be much of a "labour surplus" in the sense of an available supply of labour seeking opportunities for work and not finding an outlet; while at the same time, a big potential "labour reserve" actually exists and could be put to work and more efficient work, if appropriate policy measure were taken and could be made effective.

That a definite meaning cannot be given to a concept "underemployment" or "labour reserve" except in terms of policy holds true as a general proposition. In the concrete case of a particular country, the analysis must then be directed at ascertaining the functional and operational relationships between policies and the more or less full and intensive use of the existing, and at a future time foreseen, labour force. A study of this functional relationship, is, indeed, a crucially important analytical task in the efforts of planning for development.

In regard to the valuations as to whether, to what extent, and by what means a specific policy line should be followed so as to reach a fuller and more intensive use of the labour force, there are, as a matter of fact, very wide differences of view in all the underdeveloped countries—as also among the observers from the Western countries. Whether and to what extent the development efforts should be directed upon increasing the opportunities for work—now and in the more distant future—and increasing labour input and supply and raising labour efficiency is not self-evident. Institutions and attitudes would have to be changed through educational campaigns and/or compulsion. Vested interests in the *status quo*, firmly rooted in the existing power relations in the state and in the local communities, would have to be overcome. The entire complex of problems of land reform and tenancy reform, for instance, are involved. So are the level and distribution of taxation.

Directing, or intensifying, policy efforts in the direction of raising labour input and labour efficiency—i.e., the productivity of the labour force—implies in all directions facing issues on which views and interests are clashing. In every underdeveloped country the very diverse and conflicting political valuations in all these matters are merely concealed by the glib assertions that there is a large volume of "underemployment" and that it is a paramount interest to utilize that "labour reserve". To escape facing political issues by "objectifying" concepts and theories is an old tradition of economics and social sciences generally. But it should be clear that the study of the functional relationship between policies and the utilization of the labour force, which is the scientific task, cannot be assisted by the use of a concept "underemployment" in general, allowed to be "objectively" established in relation merely to the facts, and perhaps even measured, prior to any consideration of policy.

THE ROLE OF THE CENTRAL STATISTICAL ORGANISATION IN THE INDIAN STATISTICAL SYSTEM

By K. R. NAIR

Central Statistical Organisation, New Delhi

1. ORIGIN OF THE CENTRAL STATISTICAL ORGANISATION

1.1. The creation of the Central Statistical Organisation (CSO) by the Government of India in the Cabinet Secretariat in 1951 was indeed an important mile-stone in the development of the statistical system in India. For a considerable time prior to 1951 the statistical system in India had been, so to say, like a rudderless ship, tossed about in different directions and subject to various pulls according to the exigencies of the time, and to the needs of the situation. Consequently its growth had been of a haphazard nature and not according to any pre-conceived plan or programme. It is a well-known fact that prior to independence or rather prior to the second world war statistical organisations as they are known to-day did not exist in the Government of India or in the State Governments. It was during and after the war and particularly in the wake of independence, that the growth of the statistical system in India received a substantial impetus and resulted in the setting up of a considerable number of statistical units in the various Ministries and Departments of the Government of India and in the State Governments. The coming into existence, in a comparatively short span of time, of a large number of statistical units, sometimes with identical functions and responsibilities, also resulted in some overlapping and duplication of functions and avoidable waste of efforts and resources. Besides, there was also a tendency on the part of statistical units located in different Ministries to function in water-tight compartments without following any uniform pattern or procedure and without mutual consultation in their activities. All this pointed to the urgent need for a central coordinating body to organise the statistical system of the country, from an all-India perspective and to function as a central focal agency in statistical matters for national and international requirements.

1.2. It was in this context that the concept of the CSO took shape in the minds of the authorities of the Government of India. In 1945 an inter-departmental committee on official statistics set up by the Government of India under the chairmanship of the Economic Adviser had recommended the setting up of a Central Statistical Office as a coordinating machinery in statistical matters. Subsequently in 1948 the Standing Committee of Departmental Statisticians was also set up with Professor P. C. Mahalanobis as Chairman mainly for the purpose of exploratory and preparatory work in connection with the proposal to set up a permanent machinery for statistical coordination.

1.3. A joint meeting of the Standing Committees of Departmental Economists and Statisticians held on 10 January, 1949 recommended the formation of a Central Statistical Office in the Cabinet Secretariat. The Government of India accepted generally this recommendation, but to start with decided to set up a Statistical Unit attached to the Cabinet Secretariat which would function in an advisory capacity and would form the nucleus of the Central Statistical Organisation proposed to be set up for the purpose of coordination of statistical activities of the different Ministries of the Government of India and the State Governments and for promotion of statistical standards. Professor P. C. Mahalanobis was appointed Statistical Adviser to the Cabinet on 28 January, 1949 and was placed in charge of the Unit which started functioning from that day. Initially, for more than two years, the Central Statistical Unit was entirely staffed from and run by the Indian Statistical Institute.

1.4. In February 1951, the work relating to statistical co-ordination and statistical publications of a general nature which until then was being done in the Office of the Economic Adviser to the Government of India in the then Ministry of Commerce, was transferred to the Cabinet Secretariat along with a Statistician and his staff. The Central Statistical Organisation was set up in May, 1951 as an Attached Office of the Cabinet Secretariat and the Central Statistical Unit and the Office of the Statistician in the Cabinet Secretariat were amalgamated with this Organisation. The Administrative Intelligence Unit in the Ministry of Commerce, primary functions of which were pictorial and graphical illustration of the behaviour of statistical series collected from various Departments was also transferred to the Central Statistical Organisation.

1.5. Ever since its inception the CSO has been steadily growing in stature and importance and to-day occupies a preeminent position in the statistical system of India. Looking back on its origin and growth over the years, there is little doubt that the CSO will always owe a deep debt of gratitude to Professor P. C. Mahalanobis not only for his pioneering efforts to bring it into existence, but also for the sustained, almost parental, interest he has taken in its growth during the past twelve years of its existence. It is, therefore, in the fitness of things that on the occasion of his seventieth birthday we should not only acknowledge our deep sense of gratitude to him but also dwell at some length on the growth of the CSO and its role in the statistical system of India to-day. As an old worker of the Indian Statistical Institute during some of its formative years (1936-44) it is indeed a matter of great personal joy to me to undertake this task.

2. EVOLUTION OF THE GROWTH AND FUNCTIONS OF THE CENTRAL STATISTICAL ORGANISATION

2.1. To begin with, the following functions were assigned to the CSO: (*i*) to coordinate the statistical work of the Ministries and of other Government agencies, with a view to eliminating and preventing unnecessary duplication and reducing the overall cost to a minimum; (*ii*) to advise the Ministries and other Government agencies on statistical matters and to arrange inter-departmental discussions on statistical matters; (*iii*) to develop definitions and standards for improving national and international comparability; and to give continuing attention to the improvement of the quality of informations required by Government; (*iv*) to undertake special work as and when it arises, whether it be at the request of the Central Government Ministries or of the State Governments, including provision of facilities for training of statistical officers in day-to-day statistical work; (*v*) to prepare and publish the *Annual Statistical Abstract*, the *Monthly Abstract of Statistics*, the *Weekly Bulletin of Statistics* and other statistical publications; (*vi*) to keep continuous touch with national organisations in other countries of the world in the context of the latest developments in methodology as well as organisation; (*vii*) to serve as the channel of communication with the United Nations Statistical Office, both with regard to the observance of international conventions relating to economic statistics and provision of data required for the regular publications and for various *ad hoc* purposes; and (*viii*) to represent graphically current statistics with a view to throwing light on the developing economic situation.

2.2. Soon after the setting up of the CSO the Standing Committee of Departmental Statisticians at its meeting held in September, 1951 drew attention to the importance of maintaining close liaison between the Centre and the States in statistical matters, and in this connection, suggested that a Conference of Central and State Statisticians should be held periodically. Accordingly the first conference of Central and State Statisticians was convened by the CSO in December, 1951. Since then. the conference has become an annual feature and the secretariat assistance rendered in this connection has become one of the important functions of the CSO. After the creation of the Department of Statistics in 1961, this conference has been redesignated as the Central Technical Advisory Council on Statistics.

2.3. Another important stage in the evolution of the CSO took place in June, 1954 when, in accordance with the recommendations of the National Income Committee, the National Income Unit which was till then functioning under the Ministry of Finance was transferred to the CSO. The work concerning National Income has since become one of the major activities of the CSO.

2.4. A further expansion in the activities of the CSO took place in September, 1954 when the Planning Commission decided that a Planning Cell should be organised in the CSO to deal with statistical work relating to Planning and that this Cell should work in close collaboration with the Planning Commission, the Central Ministries and the Operational Research Unit of the Indian Statistical Institute.

2.5. Another significant step in the growth of the CSO was the creation in January 1956 of a Population Unit with the following functions: (*i*) improvement of standardisation of concepts and definitions used in connection with the population censuses and other related enquiries in India; and population projections, (*ii*) secretariat work connected with the Standing Committee on Coordination of Population and Vital Statistics, (*iii*) examination of schemes and enquiries relating to population that are referred to the CSO for technical advice from time to time, (*iv*) preparation of briefs and memoranda on census, vital statistics and demographic studies in India for national and international conferences, inter-departmental discussions etc.

2.6. In July, 1957 the Directorate of Industrial Statistics, Calcutta was transferred from the Ministry of Commerce and Industry to the Cabinet Secretariat. The Directorate's functions at that time were (*i*) to guide and coordinate the conduct of Censuses of Manufacturing Industries in various States, and tabulate, consolidate and publish the data so collected, (*ii*) to collect monthly statistics of production of selected industries of India and to publish them, and (*iii*) to compute the monthly index of industrial production in India. Subsequently the Directorate of Industrial Statistics was merged with the CSO in 1959 and has come to be known as the Industrial Statistics Wing of the CSO located in Calcutta.

2.7. Consequent on the transfer of the Directorate of National Sample Survey (NSS) from the Ministry of Finance to the Cabinet Secretariat in 1957 the Central Statistical Organisation also became the focal point in regard to the coordination and laying down of the programmes for various rounds of the NSS and exchange of information about sample surveys carried out throughout the country.

2.8. In November, 1958 the post of Director, CSO, was created for the first time and was held by Shri P. C. Mathew, I.C.S. till 1 December, 1962. Under his able stewardship during the past four years the organisation grew much in strength and vigour. To begin with, he also held simultaneously the post of Joint Secretary in the Cabinet Secretariat and later that of Additional Secretary in the new Department of Statistics set up in April, 1961.

2.9. In order to improve the statistical apparatus in the country with a view to ensuring that scientific and uniform methods of collection of statistical data are employed by the various agencies under the Central and State Governments, the Government of India decided that it was necessary to have a Central Statistical Coordinating Organisation with adequate authority to consider statistical matters; to advise and issue general directions regarding the setting up of standards and norms and methods of collection to all the Ministries, Departments of the Government of India and the State Governments; and to deal with references from the Ministries and State Governments on such questions. Accordingly, the Department of Statistics was created in the Cabinet Secretariat in April, 1961.

2.10. The Department of Statistics has an advisory body consisting of statistical experts drawn from the various Ministries of the Government of India and from the States as well. This body is designated as the Central Technical Advisory Council on Statistics which is to meet at least once a year. In addition, the Department is also assisted by a Standing Advisory Committee which is a more compact body meant to meet at more frequent intervals than the Technical Advisory Council.

2.11. Till 11 October 1960, the Central Statistical Organisation functioned as an attached office of the Cabinet Secretariat. The Secretariat Wing of the Cabinet Secretariat dealing with statistics was located at Rashtrapati Bhavan while the CSO was located in Yojana Bhavan. Since both the offices of the then Joint Secretary, Cabinet Secretariat as incharge of Statistics Wing and the Director, CSO were held by Shri P. C. Mathew, as a measure of administrative improvement, it was decided to merge the Secretariat Wing dealing with Statistics in the Cabinet Secretariat with the CSO with effect from 12 October, 1960 to function as a composite set up with location in Yojana Bhavan. This was decided upon as a measure of streamlining the administration and to speed up the work by cutting out avoidable duplication. Even after the constitution of a separate Department of Statistics in the Cabinet Secretariat with effect from 17 April 1961, the composite character of the Organisation has continued to be retained.

2.12. Mention may also be made here of the constitution of an expert committee in September 1960 by the Planning Commission under the chairmanship of Professor P. C. Mahalanobis with Shri P. C. Mathew, Director, CSO as Member-Secretary to review changes in the levels of living during the First and Second Five Year Plan periods and to enquire into recent trends in the distribution of income and wealth. It was also decided that the secretariat function of the above Committee should be performed by the CSO.

2.13. More recently in the context of the national emergency created by the Chinese aggression on our borders the Government of India decided to invest the CSO with the responsibility of studying the price trends of various commodities and for this purpose a Prices Unit has been set up.

2.14. To sum up, the present functions of the CSO may be listed under the following heads: (1) coordination of statistical work in the States and Union Territories, (2) coordination of statistical work and advice to Central Ministries in respect of (*a*) labour, employment and allied statistics, (*b*) agricultural statistics, (*c*) sample surveys, (*d*) population and demographic studies, (3) industrial, trade and shipping statistics, (4) national income estimation and public finance statistics, (5) statistical work relating to planning (6) in-service training of statistical personnel, (7) councils, committees and working groups, (8) statistical publications and graphical presentation of statistical information, (9) liaison with the United Nations Statistical Office and supply of data to international agencies and to official and non-official bodies in India and abroad. The above items of work are at present handled by the ten Divisions of the CSO mentioned below:

(1) State Statistics Division; (2) Labour Statistics Division; (3) Population Division; (4) Industry, Trade and Shipping Statistics Division; (5) Sampling and Agricultural Statistics Division; (6) National Income Division; (7) Planning Division; (8) Training Division; (9) Statistical Intelligence Division; (10) Industrial Statistics Wing, Calcutta.

3. BRIEF DESCRIPTION OF THE ACTIVITIES OF INDIVIDUAL DIVISIONS OF THE CENTRAL STATISTICAL ORGANISATION

A brief description of the activity of each of the above Divisions is given in the succeeding paragraphs.

3.1. *State statistics division.* Ever since its inception the Central Statistical Organistion has been making continuous efforts to coordinate the statistical activities in the States and to give necessary advice and assistance to the State Governments from time to time. Accordingly the State Governments were encouraged to set up and develop Statistical Bureaus in each of them with a view to facilitate the coordination of statistical activities in the States and to serve as liaison between the CSO and the State Statistical Departments. As a result, the Statistical Bureaus now exist in all the States and Union Territories of India and the vast majority of the 300 odd districts of India has a district statistical office. The development of statistical machinery in the States was further facilitated by the inclusion of statistical schemes in the State Plans during the Second and Third Plan periods and by extending Central financial assistance for the above schemes. A Working Group on 'Statistics' was formed by the Planning Commission of which CSO was the convener, to examine the statistical schemes of the States included in the Plans and to make suitable recommendations. The CSO has been convening this Working Group meetings each year and keeping a continuous watch on the progress of the statistical schemes during implementation by the States. As a result the statistical machinery in the States has expanded considerably in recent years.

Another important landmark in the coordination of statistical activities of the States and Centre was reached, as already mentioned, through the annual conferences of Central and State Statisticians which has been convened by the CSO since 1951. The conference has served as a useful forum for the statisticians of the Centre and the States for discussing important statistical problems of mutual interest and making suitable recommendations with a view to evolve an integrated statistical system in the country. Beginning with its tenth session held in December, 1961, this conference has, as already mentioned, been redesignated as the Central Technical Advisory Council on Statistics.

3.2. *Labour statistics division.* This Division deals with problems of labour and employment statistics referred to the CSO by the Ministry of Labour and Employment and other Central and State agencies from time to time, and renders necessary advice. Mention may be made of some of the more important items of work completed in recent years. (*i*) The family budget surveys in industrial centres and preparation of consumer price index numbers were examined and coordinated through the Technical Advisory Committee of Cost of Living Index Numbers. (*ii*) A brochure on Survey of Labour Force and Employment and Unemployment was prepared in consultation with the Indian Statistical Institute and concerned Central and State agencies, in order to promote uniformity in concepts, definitions and other related matters concerning survey of labour force and employment and unemployment undertaken by various agencies from time to time. (*iii*) A standard industrial classification was prepared in consultation with the Indian Statistical Institute and Central agencies with a view to bring about uniformity in the classification followed by the various organisations in this regard. (*iv*) A draft standard occupational classification was prepared in the light of comments received from the International Labour Organisation and the United Nations Statistical Office.

3.3. *Population division.* This Division was set up in 1956 to act as secretariat for the Standing Committee to coordinate population and vital statistics and to give continuing attention to the improvement of health, population and vital statistics. Subsequently in 1960-61 the Standing Committee was enlarged and reconstituted as the National Council of Population under the Chairmanship of the Home Minister, with Professor P. C. Mahalanobis and the Minister for Health as Vice-Chairmen. Simultaneously the CSO has been giving technical advice to concerned agencies on important problems relating to population and demographic studies.

3.4. *Industry, trade and shipping statistics division.* Ever since the transfer of the Directorate of Industrial Statistics from the Ministry of Commerce and Industry to the Cabinet Secretariat in 1957, the CSO has been responsible for the collection and publication of statistics concerning this important field. The Industrial Statistics Wing of the CSO, situated at Calcutta has been responsible for the annual survey of industries, the monthly statistics of the production of selected industries in India, and the computation of the monthly index of industrial production in India. The work relating to collection and compilation of shipping statistics was transferred to the CSO at the instance of the Ministry of Transport and Communications in November 1959. Accordingly the CSO took up the compilation of dead weight cargo statistics in accordance with the recommendations of the Master Committee. The work was subsequently transferred to the Ministry of Commerce and Industry in 1962. More recently the CSO initiated a proposal to conduct a survey of distributive trade in India and examined the proposal for a biannual survey of small establishments. This Division also handles the work of the Standing Committee on Business Machines Installations including until recently the price fixation of Hollerith punched cards for Government users.

3.5. *Sampling and agricultural statistics division.* The CSO has the overall responsibility for items of work to be taken up by the Directorate (or Field Wing) of the National Sample Survey as also the details of work concerning design, programme of tabulation, and of reports undertaken by the Indian Statistical Institute, which for these purposes is known as the Technical Wing of the NSS. In fulfilling some of these responsibilities the CSO is assisted by the NSS Programme Advisory Committee which meets normally twice a year. The CSO also examines and offers comments on the draft reports of the NSS, prepared by the Indian Statistical Institute before clearance is given by the Government of India for their publication. Further the CSO advises Central Ministries and State Governments on important problems of sample surveys and their draft reports from time to time. It also issues an annual publication called "Sample Surveys of Current Interest" based on material collected from Central Ministries and State Governments.

3.6. *National income division.* Ever since the transfer of the work relating to national income from the Ministry of Finance in 1954 the CSO has been responsible for compilation of the annual estimates of national income. Accordingly an Annual White Paper on National Income with the base 1948-49 has been brought out regularly. In addition, the CSO has also brought out the quick estimates of national income, and also projections of national income to assist the Planning Commission in the preparation of the Five Year Plans.

The estimates of national income which are published annually by the CSO were based on the concepts, methodology and source materials used by the National Income Committee in 1954. Since most of these source materials had undergone considerable changes, the CSO took up the reappraisal of the methodology of national income estimation, and as a result, revised estimates of national income for 1955-56 to 1959-60 at current and at 1958-59 prices were prepared. A consolidated report has also been published at the suggestion of the Planning Commission containing proposals for a revised series of national income estimates for 1955-56 to 1959-60. The CSO has in recent years also initiated action in regard to preparation of State income estimates in consultation with the State Statistical Bureaus.

Along with the national income work, the CSO has lately undertaken, studies relating to flow of funds, and prepared estimates of the capital formation in the Indian Union at current and at 1958-59 prices for the period 1948-49 to 1960-61.

3.7. *Planning division.* The CSO maintains a close liaison with the Planning Commission in regard to the statistical work relating to Planning. The Planning Divison of the CSO acts as the Statistics and Surveys Division of the Planning Commission with the Director CSO functioning as *ex-officio* Head of the latter Division under the over-all guidance of Professor P. C. Mahalanobis who is Member-in-charge of Perspective Planning, Statistics and more recently of Scientific Research in the Commission. Since its inception in 1955 the Planning Division has taken up a number of studies. Mention may be made of the growth models, depicting regional pattern and growth; estimates of parameter of two to four sectors model, better known internationally as the Mahalanobis 2/4 model, on the basis of empirical data; preparation of brochure on progress of selected projects during the Second Plan period; brochure on basic statistics relating to Indian economy; brochure on development schemes, itemisation and unit specification of physical targets; preparation of long term demand projections for important products like steel, cement, electricity and timber; changes in the distribution of income and levels of consumption; labour productivity in major census-of-manufacturing industries; a pilot study for improving statistics relating to import licence; studies on investment in building activities in private urban sector etc. The CSO also prepared progress reports on 40 selected plan projects; a study on the relationship between the capital, output, employment in selected manufacturing industries, a study of the investment in factories, a check list of cost estimates of new projects to be undertaken in manufacturing industries, mining etc.

3.8. *Training division.* In the context of the increasing need for qualified statistical personnel in the country and to make good the prevailing shortage, the CSO undertook the responsibility to organise various training courses in official statistics from year to year. The programme of training has been drawn up so as to cover all the important subject fields in official statistics viz. statistical organisation, population, agriculture, commerce, industry, labour, finance, transport, education, health, national income, planning etc. Following are the types of training courses which are at present conducted by the CSO:

(*i*) Senior Statistical Officers' Training Course: which is meant for officers already employed in the Central Ministries and State Governments. This course is organised for a period of seven weeks and consists of two parts, viz. the general part extending over a period of four weeks and specialised part covering a period of three weeks. (*ii*) Statistical Officers' Training Course: This course which is intended for junior officers of the Government is conducted in collaboration with the Indian Statistical Institute and has a duration of nine months split up into three parts. (*iii*) Training course for the students of International Statistical Education Centre, Calcutta: This is a course conducted in collaboration with the Indian Statistical Institute and is for a period of six weeks, with the object of affording training in official statistics to nationals belonging to some of the Asian countries, (*iv*) Evening course in Statistics: This is a one year course organised by the Indian Statistical Institute in collaboration with the CSO for teaching in theoretical and official statistics and is meant for Government employees and others in New Delhi who have had no previous training in statistics and who desire to get qualified in statistics. Recently this course has been supplemented by an advanced course for six months for those possessing a master's degree in mathematics. (*v*) On the job-training and short *ad hoc* training courses: In addition to the above, the CSO has also been arranging on-the-job training to individual officers of State Governments, usually for a period of six months at a time, to work on certain specific problems of current interest. Short *ad hoc* training courses have also been arranged for officers of the State Governments by the CSO from time to time, in order to give practical training and guidance to these officers in some particular field of official statistics.

3.9. *Statistical intelligence division.* As a central coordinating office, the CSO is responsible for bringing out a number of statistical publications of varying periodicity giving an all-India picture of available statistics in various fields. A separate Division of the CSO is looking after this work. In addition, this Division also maintains a Graphical Unit in which all important current statistical data having a bearing on the all-India economy and social life, are presented in the form of graphs, charts and other pictorial media. Besides, the CSO has been the main channel of communication in statistical matters between the Government of India and various international agencies such as the United Nations Statistical Office, the International Statistical Institute, the ECAFE etc., and various official and non-official agencies in India. Statistical data has been supplied by the CSO to the above agencies on a regular as well as *ad hoc* basis.

The statistical publications issued by the CSO are as follows: (1) Annual Statistical Abstract, (2) Statistical Handbook of the Indian Union, (3) Monthly Abstract of Statistics, (4) Weekly supplement to the Monthly Abstract of Statistics (published as a bilingual edition in English and Hindi), (5) Census of Manufacturing Industries, (6) Monthly Statistics of the Production of Selected Industries, (7) Annual White paper on National Income, (8) Annual Report of the Proceedings of the Conference of Central and State Statisticians, recently renamed Central Technical Advisory Council on Statistics, (9) Annual Report of the Activities of CSO, (10) Statistical System in India, (11) Indian Official Statistical Directory, (12) Sample Surveys of Current Interest, (13) Brochure on Progress of Selected Plan Projects, (14) Brochure on Basic Statistics relating to Indian Economy.

In addition to the above the CSO also brings out reports of various working groups and committees and a number of *ad hoc* technical papers containing results of special studies made in the various Divisions. Some of these papers have been published as special articles in the Monthly Abstract of Statistics.

Some of the more important items of statistical data supplied to international agencies are the following: (*i*) Monthly data for United Nations Monthly Bulletin of Statistics, (*ii*) monthly data to the ECAFE secretariat, (*iii*) quarterly data for United Nations Bulletin of Commodity Trade Statistics, (*iv*) quaterly data on index number of industrial and mineral production, (*v*) annual data for United Nations Statistical Year Book, (*vi*) annual data for United Nations Demographic Year Book etc.

3.10. *Industrial statistics wing, Calcutta.* As already mentioned in an earlier paragraph the Industrial Statistics Wing is mainly responsible for the following functions:

(*i*) To plan and coordinate the Annual Survey of Industries under the provision of the Collection of Statistics Act 1953 and rules made thereunder, and to tabulate results thereof,

(*ii*) to collect statistics of monthly production of selected industries in India and publish them,

(*iii*) to compute the monthly index of industrial production in India.

From 1959 onwards an integrated scheme designated the Annual Survey of Industries (ASI) has been introduced for the collection of basic industrial data, by combining the Census of Manufacturing Industries (CMI) and the Sample Survey of Manufacturing Industries (SSMI) on a statutory basis under the Collection of Statistics Act 1953. Under this scheme factories employing 50 or more workers with the aid of power and 100 or more workers without the aid of power were covered on a complete basis and the remaining factories employing 10 to 49 workers with the aid of power and 20 to 99 workers without the aid of power on sampling basis. The schedules canvassed, however, are the same both for the census and the sample sector. The schedule is revised to the extent feasible from time to time according to current needs of the users and in the light of difficulties experienced by the field staff. From 1961 the schedule has been elaborated to include more detailed labour and employment data.

The Monthly Statistics of the Production of Selected Industries, consists of production statistics of major industries in India on State-wise and industry-wise basis and also gives information in respect of installed capacity and stocks of different industries classified according to the International Standard Industries Classification of economic activity. Efforts are being made continuously to expand the scope and coverage of the industries from time to time in this publication.

The present Index of Industrial Production covers 201 industries with 1956 as base. In the matter of making it more broad-based and representative of the changing pattern of the industrial structure of the country, the present index represents a significant revision of the original index which covered only 88 industries and had 1951 as base.

3.11. *International conferences and delegations abroad.* Some of the officers of the CSO were deputed to represent the Government of India at various international statistical conferences held in India and abroad. To mention a few of them: (*i*) sessions of the U.N. Statistical Commission, (*ii*) sessions of the International Statistical Institute, (*iii*) conference of British Commonwealth Statisticians, (*iv*) conference of Asian Statisticians, (*v*) regional conference on research in national income and wealth for South-east Asia etc. India has been a member of the U.N. Statistical Commission from its inception in 1946 and Professor Mahalanobis had been Government of India's representative on that Commission throughout. He was Chairman of this Commission during 1954-58 and, before this, Chairman of the Sampling Sub-commission of th U.N. which did pioneering work in promoting the use of sampling techniques in various member countries.

3.12. *Councils, committees and working groups.* In the performance of its functions listed above, the CSO has been assisted by a number of councils, committees, working groups etc. from time to time. Mention may be made of the following: (1) Central Technical Advisory Council on Statistics formerly known as the Conference of Central and State Statisticians, (2) Standing Advisory Committee for the Department of Statistics, (3) National Sample Survey Programme Advisory Committee, (4) Committee on income distribution and levels of living, (5) Technical Advisory Committee on Cost of Living Index Numbers, (6) Standing Committee on Planning Statistics, (7) Working Group on Development Statistics, (8) Standing Committee on Training, (9) Indian Conference of Research in National Income, (10) Advisory Committee on National Income, (11) Working Group on Flow of Funds accounts, (12) Standing Committee for Coordination of Business Machine Installations, (13) Committee on Inter-Industry Studies, (14) Committee on Imrovement of Labour Statistics, (15) Wholesale Prices Index Revision Committee, (16) Working Group on Import Licence Statistics, (17) Working Group on Distributive Trade.

4. LOCATION AND TECHNICAL PERSONNEL OF THE CENTRAL STATISTICAL ORGANISATION

The Department of Statistics and the main office of the CSO are located along with the Planning Commission in Yojana Bhavan, Parliament Street, New Delhi whereas the Industrial Statistics Wing of the CSO is located at Calcutta. The main office of the CSO is manned by one Director, three Joint Directors, one Officer on Special Duty, two Deputy Directors, twelve Assistant Directors and twentyfour Research Officers with necessary complement of non-gazetted technical staff. The Office of the Industrial Statistics Wing, Calcutta consists of one Joint Director, one Deputy Director, one Assistant Director and five Research Officers besides the necessary non-gazetted technical staff. The present Director of the CSO has also been designated *ex-officio* Joint Secretary in the Department of Statistics of the Cabinet Secretariat.

5. THE INDIAN STATISTICAL SERVICE

Before concluding this article, I should mention the decision of the Government of India to constitute the Indian Statistical Service as per notification dated 1 November, 1961 in the Gazette of India. Professor Mahalanobis had played a significant part in this decision of the Government. My predecessor Shri P. C. Mathew brought to bear on this problem his administrative abilities and drive which acted as a catalytic agent in the formation of this service alongside a parallel Indian Economic Service. All the spade work necessary for the initial constitution of the service has been completed. The constitution of this service will mark a significant step in the building of an *esprit de corps* among the statistical personnel of the Central Government which is a consummation devoutly to be wished.

THE STRATEGY OF INDIAN PLANNING

By I. G. PATEL

Ministry of Finance, Government of India, New Delhi

1. INTRODUCTION

1.1. Any society which aims at economic progress must strive for an increase in the rate of saving and investment, a widespread diffusion of technical skills among its people and a framework of leadership that permits a reasonable degree of coherence in its efforts. These and many other aspects of national endeavour which are a necessary and integral part of the process of development cannot nevertheless be dignified to the status of a "strategy." For strategy implies essentially a deliberate choice—a choice of the point and timing and manner of attack on the problem at hand.

1.2. It is generally assumed that Indian planning, at any rate since the publication of Professor Mahalanobis's plan-frame in 1955, is based on a specific strategy. What precisely is this strategy of Indian planning? What is the particular element of choice that it embodies? What are the underlying assumptions that impart rationality to the choice? And if these assumptions are not satisfied in Indian conditions, is it possible to evolve an alternative strategy? Some of these questions are discussed in the present note in the hope that it might help clarify issues of significance in future planning.

2. THE STRATEGY

2.1. Perhaps the most frequently encountered description of the strategy of Indian planning is that it seeks to lay particular emphasis on "heavy" or "basic" or "producer goods" or "capital goods" or "machine building" industries. The reason for this emphasis most generally given is that it quickens the pace of capital formation.

2.2. An industry may be considered "heavy" either because its product is not light or because it requires heavy investment per unit of output. But we cannot assign any overriding economic significance to industries simply in terms of criteria such as the weight of their product or the scale of investment they call for. At best, if the industries we seek to emphasise happen to be "heavy," by and large, the term might help in conjuring up the right image. Similarly, industries or facilities which form the "basis" for the growth of other industries derive their significance from these "other" industries and it does not advance matters much in determining priorities to say that the "base" must come before the "superstructure." The priority or the emphasis must be in terms of the final products, although it goes without saying that whatever is essential for the production of the particular final products — be it power or transport or skilled labour or steel or other materials — must be produced in the right proportion and at the right time.

2.3. The same consideration applies to singling out "producer goods" industries for special emphasis. In common economic usage, "producer goods" are distinguished from final goods in the sense that unlike the latter which are directly consumed or invested, the former are necessary for the production of final goods. Needless to say, things overlap; what is directly consumed or invested may also in part be used for producing consumer goods and investment goods. But one cannot assign a priority to producer goods *per se* any more than one can do so in respect of basic facilities such as power or transport.

2.4. The question is not one simply of semantics, as failure to distinguish clearly between that which is sought to be emphasised and what is essentially a means to that which is desired makes for shortcomings in policy. Thus the emphasis simply on steel may well obscure the fact that production of steel for the manufacture of motor cars (rather than for the fabrication of machines) might in fact be a negation of the basic strategy.

2.5. In terms of final goods, then, one can speak of emphasis on investment goods as distinguished from consumer goods — and going a step further, on those investment goods which are intended for augmenting the capacity of the investment goods industries themselves. Assuming for a moment that the Indian strategy of economic development consists of growing emphasis on the machine-building industry (with a growing proportion of the machines devoted to the investment goods industries), what would be the rationale of such a strategy ?

3. Rationale in a closed economy

3.1. In a closed economy, a relatively greater emphasis on investment goods as distinguished from consumer goods is synonymous with a higher *ratio* of saving and investment in total national income; and if the output of investment goods is directed towards the investment goods industries themselves, the ratio of savings and investment to income would also rise at a faster rate than would be otherwise possible. In the extreme case, if all increments in output are in the nature of investment goods and all additional investment goods are reserved for the investment goods industries, barring what may be required to maintain the output of consumer goods, the transition from a low ratio of savings and investment to national income (say 6%) to the desired rate (say 20%) would be reached sooner than would otherwise be the case. If a rate of investment of 20% is deemed sufficient to yield the desired rate of growth of output as a whole (say 5 to 6%), the process or strategy may be reversed thereafter in the sense that the total output of investment goods may then be distributed between the consumer goods industries and the investment goods industries in whatever proportion that would be required to enable consumer goods output and investment goods output to grow at the same rate—keeping the ratio of both consumption and investment in national income more or less constant (say 80% and 20% respectively).

3.2. If consumer goods and investment goods industries (each taken together with their respective constellation of producer goods, basic facilities and the like) require the same amount of capital per unit of output, the total production of investment goods will need to be distributed between consumer goods industries and investment goods industries in the same proportion as is desired between consumption and investment (80 : 20). If, as is more likely, the investment goods are more capital-intensive, the investment goods industries would claim that much higher share in the allocation of investment goods (say 40% leaving 60% for the consumer goods industries).

3.3. The basic strategy we are considering and its rationale in a closed economy may be described, therefore, as follows. If a minimum growth rate of say, 5 percent per annum is desired, it would require a ratio of investment to income of, say, 20%. To maintain this rate of investment and this rate of growth would require a certain distribution of investment goods output between consumer goods and investment goods industries (say, 60 : 40). The higher the proportion of investment goods output that is reserved for the investment goods industries, the sooner would the desired ratio of investment be reached. The growth in consumption (i.e. the output of consumer goods) would necessarily be smaller in the meanwhile—and perhaps the growth in output as a whole as well if the capital requirements of the investment goods industries are larger per unit of output than those for consumer goods industries. But the loss in consumption or in output as a whole would be made good over a time, as the day when the desired rate of increase in over-all output (and in each constituent thereof) would begin to materialise, would come sooner. If the need to permit a certain increase in consumption—say at least to match the growth in population—limits the extent to which this strategy can be followed, there is, nevertheless, no escape from raising the ratio of investment goods going to the investment goods industries to a certain level, however gradually. This level is the one that should obtain when the desired ratio of investment in national income is reached and has to be maintained.

3.4. In short, given the assumption of a closed economy and given the technical relations between investment and output in different industries, there is a structure of production that is implicit in the objective of a desired rate of growth. The sooner this structure begins to be approximated, the sooner will the desired rate of growth be achieved. But the pace at which the structure can be approached will be a function of—and indeed be synonymous with—how fast the rate of saving (and investment) can be stepped up; and a faster pace of approaching what may be called the balanced structure of production might necessitate not only a quicker step-up in the rate of saving but also an acceptance of a lower rate of growth in the meanwhile.

4. AN OPEN ECONOMY

4.1. The case for investment goods industries in India, however, cannot be made to rest on the identity between investment and the output of investment goods which must prevail in a closed economy. Should not the fact that it is possible for India to take advantage of international division of labour make any difference to the kind of strategy it might follow in its planning?

4.2. If the ultimate objective is to attain a rate of investment high enough (20%) to permit a satisfactory rate of growth (5 to 6% per annum) in as short a time as possible, the structure of availabilities at India's command—sooner or later—must be the same as would obtain under the structure of production in a closed economy to which we have already referred. The difference that the possibility of international trade makes would, therefore, be in putting the same basket of consumer goods, investment goods for consumer goods industries and investment goods for investment goods industries (with their concomitant producer goods) at the disposal of India on the basis of a structure of production that would or could be different from that in a closed economy. Given the possibility of international trade, India could concentrate on the production of goods in which she has a comparative advantage and thereby achieve a somewhat higher rate of growth without having to forego that complex of availabilities which is pertinent to sustaining the tempo of further growth. International trade would thus be a liberating factor in the sense that it makes it possible to obtain that which is scarce in exchange for something which is not so scarce.

4.3. Two reasons might be given why, despite the consideration just mentioned, it would still make sense to try and erect a structure of production as if India were a closed economy. It might be argued that a country like India must have the capacity to produce all kinds of machines so that its growth would not be interrupted even when it is denied access to world markets for a time. An argument for autarky of this kind can lay little claim to rationality for it does not make much sense to determine one's day-to-day behaviour on the basis of what might be required in an ultimate emergency; and if the case is one of relative self-sufficiency in some line as a measure of insurance, it is at least arguable that there should be less emphasis on self-sufficiency in the case of things like machinery which can be run down for a few years than in regard to materials and essential consumer goods which exhaust their usefulness over a shorter span of time.

4.4. Alternatively, it might be argued that even if a realistic assessment of the possibility of international trade is made, the structure of production that would best assure the inevitable pattern of availabilities would not be materially different from the structure of production that must prevail in a closed economy. This may be for two reasons. First, the possibilities of increasing exports may be so limited that there is no escape from having to satisfy the bulk of the requirements of continuing development at a satisfactory pace from domestic production as distinguished from imports. Alternatively, it may be that while possibilities of trade are not inherently limited, they lie in the direction of a sufficiently broad-based structure of production which would correspond in broad outlines to the kind of structure that must prevail in a closed economy. In other words, the country's resource base might dictate a diversified economy which produces something of practically everything even if at each stage it would find it profitable to export some items from each category of production in return for similar items of import. Which of these two assumptions would be valid in India and what precisely are their implications? Or could it be that India's export possibilities are not only limited but are also to be found, such as they are, in the direction of an economy which is more or less fully diversified?

4.5. It could, of course, be argued that the possibilities of enlarging India's *present* exports are limited. The demand for tea or jute products around the world may not reasonably be expected to grow rapidly, whereas facilities for the manufacture of simple consumers' goods like cotton textiles might become so generally available among all countries that it would not be safe to rely on exports of such items for any length of time. It may also be that many of our present exports are particularly vulnerable in the face of growing protectionism or technical change abroad. The combination of all these factors could be shown to produce a strong presumption that, on the basis of the present range of exports, we would be faced with steadily adverse terms of trade or a slow rate of growth in the volume of exports — or both.

4.6. But arguments of the kind just presented with reference to existing exports do not warrant the conclusion that export possibilities *in general* would also be limited. Nor can one readily take it for granted that there are insuperable factors which must necessarily limit the demand or the competitive position of all existing exports. If development is a matter of more and more investment and progressively improved techniques, it would stand to reason that there would be scope for investing the growing resources in a manner which would enlarge export possibilities. There are a large number of commodities — in all fields of production—for which international demand would grow more than in others; and with ingenuity, demand could be created for many others. There is no reason to believe that a society which can successfully replace the imports of complicated products cannot nevertheless find export markets for a growing range of commodities if it sets out to do so. And it is not as if the only category of goods in which international trade is growing is that of investment goods. Fruits, processed foodstuffs, ordinary hospitality to tourists, synthetic fibres, minerals, producer goods such as fertilizers or steel, bicycles, automobiles and even cosmetics and handloom fabrics are items for which demand abroad will grow and can be stimulated. Whether conditions in world markets are such that the ratio of exports to national income must necessarily fall so that the ratio of imports to national income must also be brought down are ultimately semi-philosophical questions to which precise answers are neither possible nor necessary. The important fact is that there are no inherent limitations to exports in a developing economy so that it would be both unnecessary and wasteful to try and erect a structure of production with an accent on self-sufficiency from the outset.

4.7. Even if export possibilities are limited, what is the meaning of saying that this limitation must be overcome by trying to erect a self-sufficient economy? One may safely assume that whatever it is that limits exports will also limit the creation of a broad-based economic structure capable of delivering all the consumer goods, producer goods and investment goods necessary for rapid and continuing growth. More likely than not, lack of export opportunity or possibility will be only a reflection of the general lack of growth potential. And in any event, if export opportunities are limited, the attempt to press ahead with broad-based development will lead to balance of payments difficulties which can be overcome either by slowing down the rate of growth even when internal savings would be available or by relying on external capital. Even if external assistance is available, there would in practice be an understandable insistence on the part of both the giver and the receiver that the scope of such assistance be limited as far as possible. In other words, by the time a country comes to rely on external assistance to a significant extent, it would be called upon to exploit the possibilities of exports even if these possibilities are limited and could be exploited only at a price—i.e. by accepting less favourable terms of trade.

4.8. It might be argued that available resources should not be devoted to exploitation of export possibilities at less favourable terms of trade and should be devoted instead to activities which would ultimately save much more in imports (or earn much more by way of exports). But in practice, a strategy of this kind could be pursued with the help of external assistance only if there is considerable statesmanship and forbearance among the aid-giving countries. Even for the aid-receiving countries, it would require virtually an act of faith to forego even limited export possibilities in the present for the sake of much larger possibilities of import substitution (or of exports) at a later stage *if* external capital has to be paid back with interest. The wiser course, if export possibilities are really limited, would be not to rely on external capital and face the consequences thereof or to rely on external assistance only to the extent that it is available on terms which square with a realistic assessment of its future balance of payments.

4.9. Yet another version of the hypothesis of inherent limitation on exports may be that there is a degree of uncertainty that attaches to export promotion which does not apply to import-substitution so that the latter must be given a margin of preference in planning. Exports might be shut off by the trade policies of others, whereas what is saved on imports is secure. This is, of course, one of those propositions which gain in validity the more widely they are accepted as axiomatic. But it cannot alter one's view as to what would be more rational if it were possible.

4.10. The alternative hypothesis, viz., that while export possibilities exist and could be created and exploited, they lie essentially in the direction of establishing a broad-based industrial and economic structure, certainly fits the facts about a large continental country like India which has a diversified resource base. Given the fact that India has abundant resources of iron ore, coal, and other minerals, there is no reason why it could not, in time, be proficient in the manufacture of steel or engineering goods or machinery. Similarly, given its rich soil and low yields at present, it could add substantially to its output of many agricultural products; and the same would be true of a large variety of manufactured consumer goods. While India would export and import growing amounts of a large variety of goods, the general drift of the economy would be towards a well-diversified structure of production.

4.11. But where exactly does the line of reasoning just outlined lead us in the matter of priorities in planning? Granted that agricultural output on a wide front can be increased, it does not follow that it should be increased in the direction of substituting for imports rather than of increasing exports. Granted that steel can be competitively produced, it does not follow that exports of iron-ore should not be increased. And indeed, once steel begins to be produced in large quantity and efficiently, it does not follow that it should be better to employ the steel for producing machinery rather than exchange it directly for machinery. And when machines are produced (or indeed obtained by exchange), they may yield more if applied to consumer goods industries where export-potential (or indeed import-substitution) has reached a certain momentum rather than to developing import-substitution in new lines. At each stage of development, a choice will have to be made between consolidating and enlarging existing export opportunities and developing new export-potential, between consolidating and enlarging capacity for import substitution in existing lines and developing a similar capacity in new lines. The simple rule of going in for import-substitution all along the line cannot yield the best answer if only because it would spread the total effort too thinly. Nor can one claim any merit in merely minimising total imports over a period as there is no virtue in minimising imports if exports can be increased even more by not doing so.

4.12. At best one can argue that if a country's comparative advantage is likely to lie over a wide field and if some of these fields are such as to require a considerable time for this advantage to be developed fully, it should take action well in advance to be able to exploit this advantage. But even advance action which looks to the future involves a choice—a choice, first between immediate gain and larger future gain and, again, a choice between advance action of one kind and another. By all means, within limits, at each stage of development a country should try to keep a step ahead, so to speak, by investing in ventures that take time to fructify; but such ventures would be as much in the field of export-promotion as import-substitution—as indeed in the field of basic facilities and services which would be necessary for both. The necessity to look ahead, therefore, does not by itself imply that a structure of production more akin to that in a closed economy will happen to be the most desirable if we do look ahead. And even if this is the case, it does not necessarily follow that the best approach to the final structure that is likely to emerge is by trying to skip stages all the time rather than by advancing in uneven but measured steps in different directions at different times.

4.13. Once we embark on the high seas of international exchange and discard the shelter of the hypothesis of limited export possibilities, we cannot seek anchor in simple rules like import-substitution all along the line or the development of machine-building capacity for the investment goods industries at the fastest possible pace. The simple world of a strategy drawn from the model of a closed economy must give place to measurement and assessment of costs and benefits over time with due allowance for at least the interest one has to pay on foreign debt for buying time. The broad outlines of policy may resemble somewhat the dictates of a simple strategy; but the differences may be just as vital if not more. In actual practice, Indian planning today steers clear from *excessive preoccupation with self-sufficiency as a result of the necessity to export more.* But between these conflicting pulls, the actual course of the economy needs to be charted more precisely with reference to detailed assessment of available alternatives.

4.14. Whatever the allocation of resources that might turn out to be most appropriate from time to time, the basic necessity of increasing savings and investment and of enlarging technical skills and the organisational base would, of course, remain. But with these, we are back from the realm of strategy and choice to what is imperative and implicit in any plan of development.

SCIENCE AS A DIALECTICAL UNITY OF SYSTEM, METHOD, AND PRACTICE

By T. D. PAVLOV

Bulgarian Academy of Sciences

I

System, method, and practice are essential symptoms which enter into the composition of the concept of science. They do not exhaust the content of this concept but it does not exist without them.

Practice, in the words of Goethe, "was in the *beginning*" when man, detaching himself from his preceding anthropoid state, was in an uninterrupted and steadily complicating interaction with natural environment. Accumulating sensory, emotional, volitional, and intellectual experience in the interaction with natural environment and subsequently with social environment, man first worked out certain notions and ideas, and then categories and laws which, *taken as a system, are a subjective* (*human*) *reverberation of the objectively and actually existing order or system of the natural and social, material objects and phenomena.*

At the same time, both natural and social realities have undergone changes and development in one form or another. The systems of notions, ideas, categories, and laws have followed this development. Otherwise they would be apt to lag behind it to such an extent as to be devoid of all cognitive and practical significance to man, taken both as a social creature by virtue of his very nature and as a separate individual or personality.

The percepts, notions, ideas, categories, and laws of thinking are subject to change, though at the same time they preserve something in common which is intrinsic to every process and product of human cognition.

It is self-evident that the changes and development of notions, ideas, categories, and laws take place in the thinking human head or, more precisely, in man's consciousness which combines or ranges the elements of cognition into scientific systems. However, the changes and development of the scientific systems do not take place in an arbitrary manner, following personal whims, but in accordance with the changes and development of objectively real natural and social objects and phenomena which, at the same time, preserve certain common aspects and in the long run preserve deep in themselves the universal essence of the world—its *materiality*—which Friedrich Engels designated as the *unity* of the world.

If science does not reflect in a more or less adequate manner the objectively real natural and social things taken both in their natural system and in their further natural change and development, it would neither furnish objectively true knowledge about the world, including knowledge about society and man, nor would it serve as the basis of analyzing, prognosticating, and consequently of a scientifically based change of the world by man, socially organized and acting in a scientific manner.

II

It is not sufficient to realize that facts without theory and theory without facts do not make science. It is likewise not sufficient to differentiate between *objectively real facts* and *internally subjective facts* which are given to human knowledge only. The latter can also prove to be scientifically true, though only when they are confirmed by the objectively real facts which exist outside one's consciousness and independent from it. When there is no such confirmation at hand they remain barely "internally subjective facts" which cannot serve for a scientific explanation of the world or for changing it in a progressively expedient manner. Many of the "facts" pertaining to the mythological, religious, and pathological human consciousness are nothing but facts in parentheses. That is why human science has always assessed them in a negative manner, while human practice has not been able to use them in moving a single grain of sand in the field of objective reality.

Generalizing his practical experience, man was obliged to discover, formulate and systematize not only that which was *common* in his practice and in his natural and social environment but also that which was *particular* in them. At the same time, the common and the particular have never excluded but, on the contrary, they have presupposed the existence and active demonstration of that which is *unique* in man himself and in the surrounding objects and phenomena, and hence *unrepeatable.*

If scientific cognition meant *merely* reverberation of the general and particular, it would be difficult to imagine the process of discovering such things and natural laws which contain something *qualitatively new compared with the preceding scientific discoveries and practical achievements.*

Scientific cognition, which is a dialectical unity of scientific system, method, and practice, is likewise a dialectical unity of analysis and synthesis, of induction and deduction, of generalizing available forms and natural laws and of discovering qualitatively new ones. This is apt to be forgotten or underestimated at times, though the great thinkers and practical creators, in one way or another and to varying degrees, have not forgotten or underestimated it.

When V. I. Lenin generalized the revolutionary experience of the proletariate in Russia and in the world, he did not content himself with establishing only what is most common and what is least common (particular) in the world of imperialism at that time as well as in Russia which proved to be its weakest link. V. I. Lenin proved that the common chain of imperialism can be broken precisely at this weak link. And it was actually broken by the Great October Socialist Revolution and by the subsequent revolutions which were people's democratic in form and socialist in content.

A most typical and important feature of Lenin's scientific thinking and revolutionary practice is the fact that after establishing the most common and the less common (particular) natural laws and tasks of the October Proletarian Revolution, he brought his analysis and prognosis to the well-known assertion *that it was too early yesterday and it will be too late tomorrow, and that consequently the armed uprising must be carried out precisely on the day indicated by Lenin.*

Here we have a genuinely amazing example of profound dialectical unity between common, particular, and single, between collective and individual thought and action, between scientific system, method and practice.

The socialist revolutions which followed the Great October Socialist Revolution were bound to repeat what was common in the October Revolution, though at the same time they were all carried out under particular concrete conditions and tasks in the individual countries as well as under unrepeatable single circumstances and actions. The revolutionary experience of the whole world has shown that resorting to what is common alone each revolution would run the risk of becoming an abstract and fruitless scheme, whereas resorting to the particular and single alone each revolution would face the danger of lapsing into nationalistic and individualistic voluntarism and adventurism which doom it to futility and defeat.

Revolutionary experience and all social experience have demonstrated that the relation between what is single and what is common is of the same order with the relation between what is accidental and what is necessary, i.e. that the single and accidental are forms of existence and manifestation of the common and necessary. With logical necessity the causal explanation is supplemented by the statistical one; they are neither rendered absolute nor are they denied in a metaphysical manner.

The dialectical unity of the single and common, of the accidental and necessary, is first of all expressed in the fact that a scientific system is dead and fruitless when taken without movement and development; and that dead and fruitless is that system of concepts, categories, and laws which are not *subjective* images of *objectively real* things and phenomena but are simply purely subjective symbols and signs in man's consciousness which do not correspond to objective reality.

What is it that man seeks to know in objective reality ? In what direction and how far can man master and change this reality ? This is determined by human practice — industrial, social, scientific-experimental, etc. *It is precisely practice which is the starting point or the basis, purpose, determinant, and criterion of all genuinely scientific knowledge capable of explaining and of helping man in his efforts to change the existing natural and social reality.*

III

The process of the differentiation of sciences goes hand in hand with the process of their integration. An absolute differentiation of sciences would imply such dismemberment of scientific knowledge as would bring it in the last resort to absolute relativism, subjectivism, and mysticism. This is one aspect of the problem.

On the other hand, the absence of differentiation in sciences, including what are known as intermediate sciences as well as social and philosophical sciences, would render scientific cognition incapable of adequately expressing the reality of society and the world in its infinite variety of forms and dialectical development.

The ever deeper integration of sciences and their ever broader differentiation provide for their ever fuller co-ordination and subordination which are different at the different stages of development of general human scientific thinking, though at the same time they always contain something in common as well.

What is common is first of all expressed in the fact that all particular (or special) sciences examine and ascertain the specific structures, structural laws, and properties of their specific subjects, whereas philosophy examines the world as a whole, including social reality, atoms, space, time, movement, etc., though taken in their objectively real existence, i.e. independent of the consciousness of man and reverberating in it in a more or less adequate manner.

Herein lies the determination of genius given by V. I. Lenin to the *philosophic concept of matter*, unlike the various *particular-scientific concepts* which, as was pointed out earlier, examine and ascertain the structure, structural laws, and properties of their subjects, including human consciousness itself.

All this obliges us to examine scientific truth, taken in its full volume, both as a dialectical *process* of subjective reverberation of the objectively real and dialectically developing reality, and as a *result or product* of this process.

The determination of truth as a *process* does not turn philosophy (logic, dialectics, gnoseology) into psychology. Psychology is at hand only when we examine the process of cognition from a particular-scientific viewpoint, i.e. from the point of view of its structure, structural laws, properties, etc.

Similarly, the concept of truth as a *result* or *product* of thinking does not imply logicism which is at hand when we cease to conceive cognition as subjective reverberation of objectively real things and as a result or product of the general cognitive scientific process.

Philosophic cognition, detached from the particular-scientific one, becomes cognition which is abstract, barren, and absolutely unnecessary. On the other hand, when particular-scientific cognition is detached from philosophic cognition it degenerates into vulgar empiricism, while in the realm of practice this leads to practicism.

The dialectical unity between philosophy and the particular sciences, hence the dialectical unity between the most general philosophical scientific method and the different methods of particular sciences (having in mind that philosophy is not only a social science but a natural science as well) are likewise given in a process of uninterrupted differentiation and integration, of co-ordination and subordination, and, at the same time, in an increasingly dialectical unity with human practice, taken both as special or particular practice and as most general or ideological one, i.e. *as a practical attitude of man to the world.*

Under these circumstances the logical method of examination and exposition presupposes the historical method of examination and exposition and vice versa. Likewise, the general theoretical examinations and achievements, seemingly detached from concrete practice, prove of great significance to it at certain periods and to one degree or another. Conversely, the practical problems of what are known as applied sciences increasingly give rise to qualitatively new scientific problems in the very process of practice and of utilizing the scientific achievements available. The solution of these new problems is the duty of the respective special science, of philosophical science, as well as of the entire human science, taken as a united process of dialectical reverberation in human consciousness of natural and social realities which are united in the variety of their forms and are likewise developing in a dialectical manner.

Hence the logical inference that there should not and cannot be any underestimation of social sciences and of philosophy in favour of natural sciences and mathematics, nor vice versa, and that only when taken in their dialectical unity can they genuinely and correctly formulate and solve the great problems related to the ever fuller and more profound scientific *explanation* of the world and to its *changing* in an ever more progressively expedient manner.

In this way human science is in point of fact advancing along the road determined in a classical manner by Karl Marx in the theses on Ludwig Feuerbach : "Philosophers have merely explained the world, whereas the task is to change it."

We shall only add on this occasion that Marx, referring to philosophy, conceived it precisely as a science and that his words about philosophers and philosophy are, as a matter of principle, valid for all scientists and sciences which not only explain the world in different ways but are used and must be used by man as an indispensable means for the further progressively expedient change of the world.

IV

The fundamental and general theses of Marxism laid down by Karl Marx in his classical work entitled *Capital* were not only safeguarded from revisionist and dogmatic distortions but were further developed in a creative manner by V.I. Lenin. Why, how, and in what direction ?

Lenin ensued from the new and imperialist era of capitalism which was the era of proletarian revolutions as well.

Marx, too, had come to the idea of the significance of the alliance between workers and peasants and of the popular movement for national liberation but it was only Lenin, under the conditions of the new era and of the new practical and theoretical tasks of the revolutionary proletariate, that could elaborate on this idea and develop it further on in a creative manner. He foresaw the proletarian revolution and its subsequent stages, and set up the revolutionary proletarian party, its programme, strategy, tactics, and organization enabling it to assume the leadership of the triumphant socialist revolution in a particular country and at a particular moment, and after that to embark on the planned building of the material and technical basis of socialism and of the entire social system of socialism.

The Central Committee of the Leninist Communist Party of the U.S.S.R., headed by N. S. Khrushchev, are continuing Lenin's creative work today, advancing towards the gradual and planned building of the material and technical basis and of the entire social system of communism.

It is clear that in this instance there is not only a generalization of previous experience but that this is accompanied by the discovery of qualitatively new facts and laws.

Precisely at that moment, under the existing conditions and tasks of the socialist and communist revolutionary practice, it became necessary and possible for Lenin to formulate in his *Philosophy Notebooks* the idea which is unsurpassed in its profound impact : "*Practice stands higher than theoretical cognition as it possesses both the virtue of universality and of immediate reality.*" (V. I. Lenin : *Philosophy Notebooks*, Gospolitizdat, 1947, p. 169.)

In his work entitled *Theory of Reverberation* (all editions) the author of these lines, ensuing from these theses of Lenin and particularly from the latter's concept that truth needs a *living* and immediate link between object and consciousness, wrote : "This link, in agreement with Engels as well, is possible only *on account* of and *inasmuch as* consciousness, being a real property of the brain, is in itself a product of natural development, by virtue of which there can be no impassable abyss between it and nature. *In brief, truth needs not only universality and necessity, but also immediate reality which attributes objective and concrete character and significance to universality and necessity. In other words, truth must stand in a living, immediate, universal, and necessary link and interaction with real objects and phenomena.* Only then and to such a degree do we know their objective reality, their objective essence and shape, their universal and most profound nature or, which is the same thing, their materiality which is intrinsic to each actual object or phenomenon." (*Theory of Reverberation*, Moscow, 1949, pp. 220, 221.)

Further on, in the chapter on the unity of the logical and historical method of investigation and exposition, developing the same Leninist ideas, the author came to Engels's concept about the *self-consciousness of the world in and through man's self-consciousness*. However, this idea has not and cannot have anything in common with the "world self-consciousness" of idealists and mysticists.

"It is precisely at this highest stage of the development of the world and man," concludes the author in the chapter on the unity of theory and practice, "*that the power of logic* in the genuine son and representative of the most progressive social class becomes of necessity the *logic of creative human power*. *The deed of* logic grows into *logic of the deed*, as this is happening before our eyes today. And it is then that the "wonder" occurs :

" '*The weapon of criticism*' becomes '*the criticism of the weapon*' since the ideas, having captivated the masses, acquire an enormous material force. Then takes place the great historical leap from the realm of necessity to the realm of genuine human freedom, the transition from human prehistory to the genuine history of man, the final dethronement of all mystic "powers" and "deities", and man attains his supreme human power, truth, and beauty.

"All this really sounds as a fairy tale and looks like a genuine wonder. However, it is not a wonder or a tale, but the only possible and logically necessary general inference ensuing from the dialectical-materialistic teaching about the unity of theory and practice and about cognition as reverberation." (*Theory of Reverberation*, Moscow, 1949, pp. 261, 262.)

Today, when the author is writing these lines several decades after the first edition of *Theory of Reverberation*, he sees no grounds of repudiating what he wrote in the first edition, as he remains deeply convinced that these ideas and a number of others which were developed in the *Theory of Reverberation* are a true statement and at the same time a modest attempt at a further creative treatment of certain fundamental theses of Lenin, particularly *his idea of genius about the immediate nature of practice*. By this idea of his, Lenin brought about an enormous leap in the development of scientific dialectics, gnoseology, and logic. This leap, to be sure, has not and cannot have anything in common with any attempt at a pragmatic, subjectivistic, and mystic approach to practice and to its dialectical unity with scientific theory and method.

MEASURE, QUALITY, AND OPTIMUM SCALE

By NICHOLAS GEORGESCU-ROEGEN
Vanderbilt University, U.S.A.

It is difficult to contemplate the evolution of the economic science over the last hundred years without reaching the conclusion that its mathematization was a rather hurried job. Undoubtedly, even social sciences could not remain completely indifferent to the loud cry "science is measurement." But the haste in adopting the method supposed to elevate political economy to the rank of "true science" had some unfortunate consequences : many epistemological issues, which ought to have been clarified before any attempt at using the new tools for the old tasks were ignored or avowedly by-passed. The most important of these issues is that of the relation between quality and quantity.

However we may look at the matter, quality is our most basic concept. It is definitely prior to that of quantity, for before we can speak of a measure of A relative to B we must distinguish between A and B in some way or other. And as we do not yet have a measure of A this way can be but qualitative.

A very eloquent example of the importance of the problem of quality versus quantity is offered by Professor Chamberlin's penetrating analysis of the relationship between divisibility and efficiency of factors of production.[1] For, indeed, I believe that the whole argument of that article turns upon the manner in which the passage from quality to quantity has been handled by prominent writers on the problem of scale. There is little doubt that we much too often take for granted that the quantification of quality leaves no *qualitative residual*. The arguments advanced against Chamberlin's paper by both A. N. McLeod and F. H. Hahn[2] constitute excellent illustrations in point.

As we all know, the question of whether or not there is an optimum scale for the unit of production is far from being purely academic. In a free enterprise economy, it constitutes the basic issue in any rational policy concerning monopolistic competition; in a planned economy, it has a decisive bearing upon allocation policies. Last, but not least, it constitutes the main issue in the controversies over management by central authority.

---

[1] Edward H. Chamberlin, "Proportionality, Divisibility and Economies of Scale," *Quarterly Journal of Economics*, **lxii** (1948), pp. 229 ff. All subsequent references, however, will be made to the sixth edition of Chamberlin's *Theory of Monopolistic Competition*, which includes a more recent version of that paper as Appendix B.

[2] "Proportionality, Divisibility, and Economies of Scale : Two Comments," *Quarterly Journal of Economics*, **lxiii** (1949), pp. 128-143.

Chamberlin's thesis, therefore, is much too important to be left without an analysis by formal logic, the only one that can prevent the argument from mazing inconclusively between the multiple connotations of the same word. Such an analysis is attempted in the present paper. Its first part deals with the general problem of measurability, which has to be clarified before anything else. The second part takes up Chamberlin's problem of divisibility and proportionality versus efficiency.

To avoid at least some *ab initio* misunderstandings, some preliminary remarks appear necessary. First, I should explain that by "the passage from quality to quantity" I do not mean "the dialectical relation between quality and quantity" as taught by Hegel and understood by Hegelians of all strains. While I do not deny the existence of this philosophical problem, I proceed from the assumption—unquestionably borne out by the brute facts—that we can actually perform the specific operations required by the theoretical concept of measure. The paper being intended as a piece of *analysis* cannot enter into the deep waters of *dialectics*. The principle, to which we overtly subscribe, that quality cannot be quantified without a qualitative residual may sound as an echo of the Hegelian dogma that quantitative variations bring about qualitative changes; and actually, there are points of contact between the two. But nothing prevents us from accepting the former as a factual truth based upon the structure of the laws of the most quantitative of all sciences, physics.

Second, I should also explain that, following the teachings of Ernst Mach, by "theory" I understand a logically— as opposed to taxonomically or lexicographically —organised *description* of a definite *reality*. To the point : theorems regarding magic squares do not constitute a theory in this sense as long as they do not help us describe a definite reality. In addition to *logical* propositions, such as "If A, then B," a theory must contain also at least one *factual* proposition of the type "There exists some element in reality that corresponds to A," even if such propositions are not always explicitly listed. It is rather puzzling that economists need be continually reminded of this, while other scientists need not. Indeed, no physicist would ever regard as physical theory a piece of analysis proceeding, among others, from the proposition that light travels along a conical helix—for instance.[3]

I. MEASURABILITY AND MEASURE

1. *General considerations.* The problem of measurability of utility has by now been distilled and redistilled. In contrast with this situation, the measurability of the elements involved in production has received—to my knowledge—no systematic treatment. The anachronism is probably due to the fact that the protests against the mathematization of political economy have been exclusively directed against the idea of submitting human actions to mathematical rigor. Naturally, mathematical economists have been wont to strengthen precisely that position against which their opponents concentrated their attack. Speaking in retrospect, one should have expected the problem of measurability to be more intricate in production than in

---

[3] The reader will have no difficulty in seeing that in their criticism both McLeod and Hahn (*op.cit.*) adopted a different meaning of "theory" than the Machian one.

consumption theory. The comparability of utility has a solid basis in the concept of "the *same* individual," who chooses without ever changing his mind. Moreover, in the case of utility measurability concerns only one element: all others, the commodities involved, are assumed to possess a definite measure — an assumption generally though not absolutely true to facts. Both these simplifying conditions are not present in the productive sector where we need a measure for both products and factors and where comparability must rest on an objective instead of a subjective basis. It is, therefore, clear that even the simple concept of production function, as this is used throughout the literature, is far more delicate than that of ophelimity function. The contrast between tho two is accentuated by the fact that, while the most important propositions of the theory of consumer's behaviour are unaffected by whether utility is ordinally measurable or only ordered, there are no orientation marks for the behaviour of the producer unless factors and products are measurable in a particular sense.

By now, we are familiar with the most general concept of measure, i.e., the ordinal measure. This rests on the following definitions:

Definition 1: A set, S, is said to be ordered by a relation, R, if for $a, b, c \in S$,

(i) either aRb or bRa;
(ii) aRb and bRa implies $a = b$;
(iii) aRb and bRc implies aRc.

Relation R is referred to as an ordering relation of S.

Definition 2: A set, S, ordered by a relation, R, is ordinally measurable if there exists a numerical function, $f(x)$, $x \in S$, such that for $a, b \in S$

$$f(a) \geqslant f(b) \Longleftrightarrow aRb.$$

Any function $f(x)$ satisfying this condition is an ordinal measure of S. It is elementary that if $f(x)$ is an ordinal measure of S and $F(t)$ an increasing function of t, then $F(f)$, too, is an ordinal measure of S. Or, as this is frequently stated, ordinal measurability is indifferent to monotonic transformations of the scale.

The usual conclusion reached by most studies of measurability by both economists and psychologists seems to be that all measures are ordinal and there is no reason whatever for attaching any special significance to the concept of cardinal measures. According to this view the old controversy between cardinalists and ordinalists turns around a bogus. However, these studies fail to take into account the manifold aspects of reality. While it is perfectly true that any usual scale can be converted into the most bizarre one by a monotonic transformation, it is far from certain that we could always live by the latter without first reconverting it into the former. And if in some cases we must reconvert, the proposition that any scale is as good as any other becomes a purely analytical statement without any significance as far as scientific theory is concerned.

For a psychologist as well as for an educator it does not matter at all whether certain ability is graded on a scale from 1 to 4 or from -2 to $+2$, or even on an alphabetical scale. One may speculate about a world in which a carpenter says "two" after laying down his yardstick four times, and "three" after laying it down nine times.

But it is inconceivable that in such a world money should be counted in an arbitrarily different manner. That is not all. If we maintain that any scale is as good as any other, then such fundamental notions as *decreasing marginal rate of substitution*, *constant returns*, *efficiency*, etc., loose any meaning whatsoever. A most simple example will show this clearly.

Example 1 : The production function $q = \sqrt{xy}$ exhibits constant average returns to scale as well as decreasing marginal rate of substitution. Let us adopt a new scale for x obtained by the "smooth" monotonic transformation

$$(T) \quad x = e^{-2/u} \quad \text{for} \quad 0 \leqslant x \leqslant e^{-2}, \quad x = e^{2u^2-3} \quad \text{for} \quad e^{-2} \leqslant x.$$

If we use the same transformation for changing y into v, the production function in the domain $u, v \geqslant 1$ is given by $q = e^{u^2+v^2-3}$. In that domain (at least) the new isoquants, $u^2+v^2 = \text{const.}$, are concave towards the origin. Constant average returns, too, have disappeared from the picture. Which ordinal scale shall we then consider as proper for describing the essential properties of the production function? Most certainly, ordinal measurability does not fit even the simplest picture of a production process.

Nor can we follow the advice offered by ordinalists and choose that scale for which the formula of the given production function would be as simple as possible. For the same factors enter into a host of other production functions, and the ordinal scale that would simplify one particular production function is apt to render others forbiddingly complex. We are not in the same position as the physicists; they were able to choose, for instance, a particular ordinal scale for temperature in order to obtain the simplest possible expression for the basic equation of thermodynamics. On the contrary, we are bound to some specific procedures of measurement that are applied to commodities without number, and our task is to find out a notion of measure that would bring to the surface the common essence of all these procedures. For our behaviour as producers (or planners) has been shaped on that basis over a very long time and, consequently, cannot be separated from it. The point is that these procedures could not have resisted the test of time if they, in turn, were not rooted in some basic property of matter. And this property is that of being *cardinally* measurable.

2. *Cardinal measurability*. Instead of writing down at once all the axioms that a cardinally measurable set, C, fulfils, it is preferable to introduce them one by one. In this way we shall be able to discuss every axiom from two viewpoints : epistemological, in order to establish its status as an *operational* description of reality, and logical, in order to test the theoretical construction of the final concept.

One of the most conspicuous facts of everyday life is that we can perform a special kind of operation with some variables but not with others. We can, for instance, pour one cup of water and one glass of water into a pitcher : the result is another concrete instance of the same logical category, "water". Clearly, in other cases this operation is not feasible. We cannot always subsume two words into another word. There is no sense in which utility, calendar dates, or temperature-levels, can be subsumed. However, we can subsume two colours by mixing two pints of liquids having those colours. Another case of subsumption will interest us later on : if a bicycle frame and a set of bicycle tires are both products, then a bicycle, too, is a product. This subsumption corresponds to the fact that production processes can be consolidated (cf. Sec. 3, B4). What counts is that *some operation exists by which subsumption is carried out.* Moreover, in all cases where such an operation exists, its result is not affected by the order in which the elements involved are manipulated. We are thus justified in introducing the following

Definition 3 : An internal operation, $\perp$, of a set, S, such that

(i) if $a, b \in S$, then $a \perp b = b \perp a = c, c \in S$,

(ii) if $a, b, c \in S$, then $(a \perp b) \perp c = a \perp (b \perp c)$,

is called subsumption.

It is also characteristic of many variables in the real world that "nothing" has an unambiguous meaning in connection with them, i.e., a particular concrete instance corresponds to the basic notion of "nothing". (The latter should not be confused with the number *zero* which—as illustrated by the ordinary temperature scale—does not necessarily mean "nothing".) One property of "nothing" is reflected by

Definition 4 : If there is an element $z \in S$ such that if $a \in S$, then $a \perp z = a$, z is a neutral element.

It is instructive to observe that though colours can be subsumed by mixing, the set of all colours has no neutral element. There is no colour that can be mixed with *any* colour without changing the latter.

Definition 5 : A set that possesses a subsumption and a neutral element with respect to it, is a subsumptive set.

As the first property of a cardinally measurable set, we shall list

A1 : The set, C, is subsumptive.

This property does not permit the analysis to be carried very far. But we can easily prove the following

Theorem 1 : *The set C has only one neutral element.*[4]

We may also introduce a relation, G = "greater", through

Definition 6: If $a = b \perp c$, and $c \neq z$, then $a \, G \, b$.

At this point one may be tempted to believe that C is ordered by G. Unfortunately, proposition A1 does not exclude the possibility that both $q \, G \, p$ and $p \, G \, q$ should be true for some $p \neq q$. (See Example 2 below.)

The next three propositions pertain to the following abstract problem :

If $a, b \in C$ *does the equation* $a \perp x = b$ *have a solution* $x \in C$?

---

[4] This is elementary : if z and y are neutral elements, then by Definition 4, $y = y \perp z$ and $z = z \perp y$, and by Definition 3, $y \perp z = z \perp y$; hence $y = z$.

Or simpler, can the elements of C be *subtracted* as well ? We know that we can take goods out of a storehouse or a pail of water from a lake, so that the problem raises unavoidably. Let us begin with the simplest properties :

A2 : If $a, b \epsilon C$ and $a \perp b = a$, then $b = z$.

A3 : If a, $b \epsilon C$ and $a \perp b = z$, then $a = b = z$.

Intuitively these are so obvious properties of "nothing" that one may be inclined to believe them to be direct consequences of A1. The following example proves that they are not.

Example 2 : Let S be the set of all ordered pairs (m, n) where m, n are real numbers and $m \geqslant 0$. Let the operation $\perp$ be defined by

$$(m, n) \perp (m', n') = [\max(m, m'), n+n'], \qquad \ldots \quad (1)$$

where max (m, m') stands for the greater of m and m'. It is immediate that S satisfies A1 with $z = (0, 0)$. Yet, according to (1),

$$(m, n) \perp (m, 0) = (m, n),$$

which for $m > 0$ shows that S does not satisfy A2. Also

$$(0, n)+(0, -n) = (0, 0), \qquad \ldots \quad (2)$$

which for $n \neq 0$ shows that S does not satisfy A3.

We can use this example to prove a point made earlier, namely, that A1 does not warrant the irreflexiveness of G. Indeed, if $n \neq n'$ we have both (m, n) $G(m, n')$ and (m, n') $G(m, n)$, for $(m, n) = (m, n') \perp (m, n-n')$ and $(m, n') = (m, n) \perp$ $(m, n'-n)$.

The following two examples are designed to prove that neither A3 can be derived from A1-A2, nor A2 from A1 and A3, so that A1, A2 and A3 are not redundant.

Example 3 : Let S' be a subsumptive set satisfying A1, and let y be its neutral element. Let $z \notin S'$, and let S be formed from S' by the inclusion of z. Let us also put $a \perp z = z \perp a = a$, for any $a \epsilon S$. It is easy to prove that S is subsumptive and satisfies A1 and A3, but not A2. Indeed, if $b \epsilon S'$ we have $b \perp y = b$, although (by Theorem 1) $y \neq z$.

The example also proves the following

Theorem 2 : *The set obtained by eliminating the neutral element from a set satisfying* A1 *and* A2, *has no neutral element.*

The theorem shows that A2 expresses the absoluteness of the notion of "nothing"

Example 4 : Let $S = \{a, b, c\}$ and let $\perp$ be defined by the table

	a	b	c
a	a	b	c
b	b	c	a
c	c	a	b

The structure thus defined satisfies A1 and A2 with $z = a$, but not A3. Indeed, $b \perp c = z$, although $b, c \neq z$.

This example also shows that A3 expresses the fact "nothing" cannot be greater than any other "something" : from $b \perp c = z$, it follows that zGb.

We need also to prove that A1-A3 are not inconsistent. That can be done by referring to the set of non-negative numbers with $\perp$ defined by the usual addition. However, an example from the familiar field of economics will prove more instructive.

Example 5 : Let S be the set of all ordered pairs (m, n), where m and n are non-negative numbers. Let $\perp$ be defined by

$$(m, n) \perp (m', n') = (m+m', \; n+n'). \qquad \ldots \quad (3)$$

It is obvious that this set satisfies A1-A3. Moreover, in this case relation G is irreflexive, for $(m, n)\, G(m', n')$ implies $mn > m'n'$, $m \geqslant m'$, $n \geqslant n'$. Therefore we cannot have also $(m', n')\, G(m, n)$. In addition, G is transitive as well. Yet, S is not ordered by G; for if $m > m'$ and $n < n'$, neither of the equations $(m, n) \perp (x, y) = (m', n')$, $(m', n') \perp (x, y) = (m, n)$, has a solution in S. The solution of the first equation is $(m'-m, n'-n)$, of the second, $(m-m', n-n')$; they do not belong to S since $m'-m < 0$ and $n-n' < 0$. As we usually put it, the set S cannot be ordered because not all of its elements are *comparable* through the only available relation, G. The purpose of the "preference" concept in consumer's theory is precisely to introduce a more inclusive relation than G, so as to render any two elements comparable. In the case of a cardinally measurable variable, comparability derives from the *homogeneity* of the elements covered by it. The latter property is defined as follows :

Definition 7 : A subsumptive set, S, is homogeneous if for any $a, b \in S$, at least one of the equations

$$a \perp x = b, \qquad b \perp y = a, \qquad \ldots \quad (4)$$

has a solution.

Equations (4) may have a solution for some elements even though the corresponding set is not homogeneous. However we can prove

Theorem 3 : *If S is a set having the properties* A1-A3, *and if both equations* (4) *are satisfied for given $a, b \in S$, then $a = b$.*

Indeed, if (4) are both satisfied we have $a \perp x \perp y = a$. By A2, $x \perp y = z$, and by A3, $x = y = z$. By Definition 4, $a = b$.

A4. The set, C, is homogeneous.

We may note that Example 5 shows that a set fulfilling A1-A3 does not necessarily fulfil A4; indeed, if $m > m'$ and $n < n'$, neither of the equations (4) has a solution for $a = (m, n)$, $b = (m', n')$. On the other hand, the set of Example 2 satisfies A1 and A4, but not A2 and A3; this proves that neither A2 nor A3 follows from A1 and A4.

Theorem 4 : *If S is a set having the properties* A1-A4, *and if, given $a, b \in S$, the equation $a \perp x = b$ has a solution, then this solution is unique.*

Let us assume that we have $a \perp x = b$ and $a \perp y = b$; by A4 we have either $y = x \perp t$ or $x = y \perp v$. Because of the symmetry present, we need consider only the first alternative. We have $a \perp x \perp t = b \perp t = b$. By A2, $t = z$; hence $y = x$.

Definition 8: If $a, b \in S$, and $a \perp x = b$ has a solution $x \in S$, then x is the difference between b and a, and we write $x = b \top a$.

From Theorems 3 and 4 we derive immediately

Theorem 5 : *If a set* S *has the properties* A1-A4, *and* $a, b \in S$, *then there is only one difference connected with the pair* (a, b) *that is, only one of the following alternatives are possible* :

(i) $a \neq b,\ (b \top a) \in S,\ (a \top b) \notin S,$

(ii) $a \neq b,\ (a \top b) \in S,\ (b \top a) \notin S,$

(iii) $a = b,\ (a \top b) = (b \top a) = z.$

From Definition 6, it follows that if we have (i), then bGa, and if (ii), then aGb. If we now introduce

Definition 9 : Relation $a\Gamma b$ means either $a = b$ or aGb,
we can easily prove the following

Theorem 6 : *A set having the properties* A1-A4 *is ordered by* Γ.

This result shows that A4 is a truly critical step in our argument. We can subsume "apples" and "pears" into "baskets of apples and pears." But subsumption alone does not guarantee that we can always perform the inverse operation, that of subtracting one element from another. No one can subtract a "basket of two apples and five pears" from a "basket of four apples and three pears," or vice versa. The reason is that while "apples" and "pears" are, each by itself, homogeneous substances, "baskets" are not so. *It is subtraction, not subsumption, which provides the acid test for homogeneity.* A4, therefore, can be properly called the *Homogeneity Axiom.*

Now that we know that a set fulfilling A1-A4 is ordered, the question arises of whether such a set lends itself to some sort of measure. As the following example shows, not every set which satisfies A1-A4 is ordinally measurable. (On the other hand, it is hardly necessary to point out that not every ordered set satisfies those same conditions.)

Example 6 : Let S be the set of all ordered pairs (m, n), where m and n are real numbers such that either $m > 0$ or $m = 0, n \geqslant 0$. Let operation $\perp$ be defined by (3). It is immediate that S has the properties A1-A4 and, hence, is an ordered set. The order is lexicographic :

$$\begin{aligned} &(m, n)\Gamma(m', n') \quad \text{if} \quad m > m', \\ &(m, n)\Gamma(m, n') \quad \text{if} \quad n \geqslant n'. \end{aligned} \qquad \ldots \quad (5)$$

However, we can prove the following

Theorem 7: *There is no ordinal measure of* S *based upon relation* Γ.

Let us suppose that there is an ordinal measure of S and let us write $f(x, y)$ for the measure of (x, y). Considered as a function of y, $f(x, y)$ is an increasing function; moreover, for any $x < +\infty$, we have $-\infty < f(x, y) < +\infty$. Therefore, for $x > 0$ we can put

$$h(x) = \lim_{y \to -\infty} f(x, y), \quad H(x) = \lim_{y \to +\infty} f(x, y). \qquad \ldots \quad (6)$$

Obviously,
$$G(x) = H(x) - h(x) > 0$$
for $x > 0$. Let $I(\alpha)$ denote the set of all values of x such that (1) $0 < m < x < M$ and (2) if $x \epsilon I(\alpha)$ then $G(x) > \alpha > 0$. There exists a δ for which $I(\delta)$ is an infinite set. Indeed, let us presuppose that $I(\alpha)$ is finite for any $\alpha > 0$, and let us consider a positive sequence $\{\alpha_i\}$ such that $\alpha_i > \alpha_{i+1}$ and $\alpha_i \rightarrow 0$ for $i \rightarrow \infty$. Every value x of the open interval (m, M) would then belong to an $I(\alpha_i)$ as well as to all $I(\alpha_j), j > i$. In this case, all values of the open interval (m, M) can be arranged in a sequence $(x_1, x_2, \dots x_{a_1}, \dots x_{a_2}, \dots)$. But this is absurd, for the set represented by (m, M) is not numerable. Hence, there exists an infinite set $I(\delta)$. This set contains a strictly monotonic sequence $J = \{m_i\}$. (We need consider here only the case where $m_i < m_{i+1}$.)

Let us now choose $\varepsilon > 0$ such that $\eta = \delta - 2\varepsilon > 0$. Because of (6), for every $m_i \epsilon J$, there is an N_i and an n_i such that
$$f(m_i, N_i) > H(m_i) - \varepsilon, \quad f(m_i, n_i) < h(m_i) + \varepsilon.$$
This yields
$$f(m_i, N_i) - f(m_i, n_i) > G(m_i) - 2\varepsilon > \eta.$$
Hence, $N_i > n_i$, and by (5) and Definition 2,
$$f(m, n) < f(m_1, n_1) < f(m_1, N_1) < \dots < f(m_k, n_k) < f(m_k, N_k) < \dots < f(M, N).$$
Therefore,
$$f(M, N) - f(m, n) > \sum_{1}^{k}{}_i \, [f(m_i, N_i) - f(m_i, n_i)] > k\eta,$$
for any k. Since this inequality is obviously impossible for $k \rightarrow \infty$, the theorem is proved.

This result shows that we have not yet taken into account all properties connected with cardinal measure. The missing property is connected with the name of Archimedes who first pointed out its importance for the concept of measure.

A5 : The set C possesses the Archimedean property, i.e., given $a, b \epsilon C$ and $a \neq z$, there exists a positive integer k such that $(ka)\, G b$.[5]

The property described by A5 corresponds to the transparent fact that if the water in a reservoir, for instance, is to be measured with the aid of a pail, we must be able to empty the reservoir by removing a finite number of pails of water. It is clear that the set of Example 6 does not possess the Archimedean property : if $m > 0$, then $(m, n)\, G[k(0, 1)]$ for any k. But that set is not ordinally measurable. On the other hand, we can prove that the Archimedean property does not necessarily guarantee that an ordered set should be ordinally measurable.

Example 7 : Let S be the set of all ordered pairs (m, n), where $m > 0$ and n are real numbers, and let $\perp$ be defined by (3). The set is obviously ordered by the relation R :
$$(m, n)\, R(m', n') \quad \text{if} \quad m > m',$$
$$(m, n)\, R(m, n') \quad \text{if} \quad n \geqslant n'.$$

---

[5] By (ka) we denote the element $a \perp a \perp \dots \perp a$ where $\perp$ is applied $n-1$ times.

Moreover, S possesses the Archimedean property for if $(m, n), (m', n') \epsilon S$, there obviously exists a k such that $(km', kn')\ R\,(m, n)$. Yet, S is not ordinally measurable (via Theorem 7). Let us observe, however, that in this case there is no strict relationship between R and $\perp$, because S does not satisfy A4, i.e., S is not homogeneous.

We can now reach the following important result :

Theorem 8 : *A set, C, possessing the properties* A1-A5, *has an ordinal measure,* meas (x), *such that if $a, b \epsilon S$ then*

$$\text{meas}\ \ (a \perp b) = \text{meas}\ (a) + \text{meas}\ (b). \qquad \dots \quad (7)$$

The proof of this theorem follows very familiar lines. An arbitrary $u \epsilon C$, $u \neq z$ is chosen as unit, i.e., meas $(u) = 1$. Then, for any $v \epsilon C$, $v \neq u$, a Dedekind cut, r, of the non-negative rational numbers, m/n, is obtained according to whether $(mu)\ G(nv)$ or $(nv)\Gamma\,(mu)$. It is a simple matter to show that by putting meas $(v) = r$ we satisfy the theorem. Clearly, meas $(z) = 0$.

Definition 10 : A set, C, satisfying A1-A5 is cardinally measurable and any measure satisfying (7) is a cardinal measure.

Theorem 8 calls for a few remarks. First, if a cardinally measurable set, S, contains an element, u, such that uGx, $x \epsilon S$, implies $x = z$, it follows that if $a \epsilon S$, then $a = (nu)$. Indeed, if $a \epsilon S$, and $a \neq (nu)$, then by A5, we have $[(n+1)u]\ GaG(nu)$, and by A4, $a \top (nu) = u_1 \epsilon S$. It is immediate that uGu_1 and $u_1 \neq z$, which is contrary to our premise. On the other hand, it follows that if this premise is true, S has only one cardinal measure that does not need some small print, namely, that for which meas $(nu) = n$. There are, then, structures that provide a *raison d'être* for the concept of *natural unit.*

Secondly, if meas $(b) = k$ meas (a), we may find convenient to write $b = (ka)$, even if k is not an integer. We must, however, be careful so as not to push the parallelism too far. Indeed, while it is true that if $a \epsilon S$ then $(ka) \epsilon S$ for any non-negative integer, this may no longer be true if k is not an integer; (ka) may have no meaning at all. Only in a metaphorical way can one say that ($\frac{1}{2}$ man) is an element of the set "men". This remark leads us to introduce

Definition 11 : A cardinally measurable set, C, such that if $a \epsilon C$, then $(ka) \epsilon C$ for any $k \geqslant 0$ is a perfectly divisible set. (Undoubtedly, this definition corresponds to the perfect divisibility one encounters in all discussions of production factors.)

Thirdly, from the fact that a cardinally measurable set, S has no natural unit it does not follow that S is perfectly divisible.

Example 8 : Let S be defined as follows : there exists a $u \epsilon S$ such that if $v \epsilon S$, then $(mu) = (3^n v)$, where m and n are non-negative integers, and conversely. The set has a cardinal measure, meas $(v) = m/3^n$, and has no natural unit; but there is no $x \epsilon S$ such that $(2x) = u$, for instance. However, any cardinally measurable set has the following interesting property:

Theorem 9 : *If $a, b \epsilon C$ and $b = (pa)$, where $p = m/n$ is an irreducible fraction, then $(qa) \epsilon C$ for any irreducible fraction $q = l/n$.*

The case $m > n$ can be reduced to that where $m < n$, by observing that if $m = kn+n'$, $n' < n$, then $b = (ka)\perp b'$, aGb'. Now, if $1 < m < n$, by A4 and Theorem 8, there exists $b' \epsilon C$, such that meas $(b') = m'/n$, $m' = n-m$. If $n-m = 1$, the theorem is proved. If $n-m > 1$, let $m_1 = \min\,[m, n-m]$, and let b_1 be the element having m_1/n as measure. By A5, there exists an integer k such that $[(k+1)b_1]\,GaG(kb_1)$. Clearly, we cannot have $(kb_1) = a$, for that would mean either $km = n$ or $k(n-m) = m$, and this is impossible because m and n are mutually prime. Therefore, $a \top (kb_1) = b_2 \neq z$. And since $b_1 G b_2$, it follows that meas $(b_2) = m_2/n$, with $0 < m_2 < m_1$. If $m_2 = 1$, the theorem is proved; if $m_2 > 1$, we continue the algorithm, obtaining $m_1 > m_2 > \ldots > m_k$, with meas $(b_k) = m_k/n$. The algorithm inevitably stops when $m_k = 1$, which proves the theorem.

Fourthly, an ordinally measurable set is not, necessarily, cardinally measurable. Cardinality requires a structure with respect to which subsumption *and* subtraction should have a definite operational meaning distinct from the arithmetic manipulation of the ordinal scale. This point has an important bearing upon the prevalent view (originated by Pareto) that comparability of utility differences leads to a meaningful utility scale which leaves only the origin and the scale unit arbitrary.[6] The fact is that this latter type of measurability involves a stronger condition than comparability of differences, namely, their cardinal measurability. Undoubtedly, either this latter condition has been smuggled, as it were, in the arguments concerning the measurability of utility, or the results of such arguments have been misinterpreted. The following will make this point clear.

Definition 12 : A set, S, has a weak cardinal measure if the following conditions are fulfilled :

(i) There exists a cardinally measurable set, Σ, associated with S through an external operation, $\overset{0}{-}$, such that if $a \epsilon S$ and $\alpha \epsilon \Sigma$, then $(a \overset{0}{-} \alpha) \epsilon S$.

(ii) For any $a \epsilon S$ and $\alpha, \beta \epsilon \Sigma$, we have $(a \overset{0}{-} \alpha) \overset{0}{-} \beta = a \overset{0}{-} (\alpha \perp \beta)$.

(iii) If $a, b \epsilon S$ there exists an $\alpha \epsilon \Sigma$ such that either $a \overset{0}{-} \alpha = b$ or $b \overset{0}{-} \alpha = a$.

(iv) If $a \overset{0}{-} \alpha = a$, then $\alpha = \zeta$, where ζ is the neutral element of Σ.

It can be easily proved (1) that S is ordered by the relation R defined by bRa if $b = a \overset{0}{-} \alpha$, and (2) that S possesses an ordinal measure, meas (x), $x \epsilon S$, such that, $a \epsilon S$ being arbitrarily chosen,

$$\text{meas}\,(a) = 0 \quad \text{and} \quad \text{meas}\,(b) = \pm \text{meas}\,(\alpha), \qquad \ldots \quad (8)$$

where meas (y) is a cardinal measure of Σ and the sign is set according to whether $b = a \overset{0}{-} \alpha$ or $a = b \overset{0}{-} \alpha$. By writing $w = x \underset{0}{-} y$ if $x = y \overset{0}{-} w$, we obtain

$$\text{meas}\,(x \underset{0}{-} y) = \text{meas}\,(x) - \text{meas}\,(y), \qquad \ldots \quad (9)$$

which characterizes the weak cardinal measure.[7]

---

[6] See, for instance, O. Lange, "The Determinateness of the Utility Function," *Review of Economic Studies*, i(1933/4), pp. 218-225, as well as the comments of E. H. Phelps Brown, H. Bernardelli and Lange's reply in *ibid.*, ii(1934/5), pp. 66-77.

[7] Since the proofs of the above results present no difficulty, there is no need for spelling them out here.

It is vitally important to bear in mind that $\overset{0}{-}$ is a distinct operation from $\perp$: for if we fail to do so, we are apt to miss the point that S and Σ do not pertain to the same "substance" and that the operation $\overset{0}{-}$ precedes in *logical and operational* order the "arithmetical" difference of scale readings which appears on the right hand side of (9). No other concrete example can illustrate the points just made as well as chronological time. The latter's weak cardinal scale is not derived by adding two elements of the same substance : the elements added are neither both "chronological dates" nor both "time-intervals," but one is a "chronological date" and the other a "time-interval". The same is true of the temperature scale—another favourite example of the discussions of measurability. To a given "temperature level" we add a certain *amount* of "energy" to obtain another "temperature level." Both illustrations plainly show that the "difference" element in a weak cardinal measure must be cardinally measurable : its ordinal measurability, so much the less comparability alone, does not suffice. One may even say in general that the reason why a particular ordinal scale is singled out as having a definite operational meaning is that the scale is related in some manner to a cardinal scale. This relation is the simplest in the case of weak cardinality.

II. DIVISIBILITY, PROPORTIONALITY AND RETURNS

3. *Individual product and individual factor.* The very fact that we have an operational, though primary, notion of a process as a transformation of some elements into others, means that we can distinguish in a very objective way between the beginning, the *input,* and the end, the *output.* These obvious remarks justify us in laying down the following "axioms" as a first description of the productive sector :

B1 : There exists a set, Π, such that if $\pi \epsilon \Pi$, then π is a concrete output and conversely.

B2 : There exists a set ϕ, such that if $\varphi \epsilon \phi$, then φ is a concrete input, and conversely.

B3 : To every $\pi \epsilon \Pi$ there corresponds at least one $\varphi \epsilon \phi$, and to every $\varphi \epsilon \phi$ there corresponds at least one $\pi \epsilon \Pi$.

The last proposition merely expresses the fact that a concrete output is produced and a concrete input is productive. We should explain that "concrete" is not used in the sense of "actual". "Concrete output" simply means one instance of the concept of "output".

Definition 13 : A pair (π, φ) of the correspondence B3 is a concrete process.

Propositions B1–B3 in no way restrict the size or the structure of the elements π and φ: a concrete output may very well be a multi-product in the usual sense of this term. On the other hand a certain *technological horizon* must be assumed for their operationality. "Brown-and-serve rolls" became a member of Π only after we discovered how to produce them, and no φ contained a "pilot" before the invention of the airplane.

B4 : Given two concrete processes, (π', φ'), (π'', φ''), there exists a concrete process (π, φ) such that $\pi = \pi' \perp \pi''$, $\varphi = \varphi' \perp \varphi''$, where $\perp$ is a subsumptive operation.[8] This proposition states that any two concrete processes can conceivably be carried out independently of each other and at the same time. The famous objection of Pareto that "there are not two Paris cities" for Paris to be used simultaneously in two distinct processes is of no avail. Material scarcity affects only our choice, not the logical relationships of our concepts : the fact that we cannot actually draw a circle larger than the equator does not limit the properties of "circle" to that size. The following three properties hardly need any commentary :

B5 : The sets Π and ϕ have a common element, z, such that if (π, z) is a concrete process, then $\pi = z$; and if (z, φ) is a concrete process, then $\varphi = z$.

B6 : The element z is neutral for Π and for ϕ.[9]

B7 : The sets Π and ϕ satisfy A2 and A3.[10]

Obviously, it would be absurd to attribute to either Π or ϕ the homogeneity property, A4. On the other hand, we must recognize that Π (and ϕ) can be divided into cardinally measurable subsets, Π_i (and ϕ_i) having no common element other than z. The idea is that "this" particular tube of tooth paste—call it D—is a concrete instance of the cardinally measurable set $\Pi_i = \{n D\}$, where n is a non-negative integer. The same is true of the concrete input which includes the particular management unit of Y Inc., irrespective of whether the latter consists of a single person or a board with a sizable staff. In this way quality is clearly separated from quantity: the qualitative difference between concrete outputs or between concrete inputs is shown by the subscript of the subsets Π_i and ϕ_i to which they belong. And since qualities may be more numerous than numbers themselves, the subscripts are not necessarily numerical.

These remarks lead to

B8 : If $\pi \epsilon \Pi$ (or $\varphi \epsilon \phi$), there exists a cardinally measurable subset $\Pi' \subset \Pi$ (or $\phi' \subset \phi$) such that $\pi \epsilon \Pi'$ (or $\varphi \epsilon \phi'$).

B9 : If $\pi \epsilon \Pi$ and $\pi \neq z$, there exists a cardinally measurable subset $\Pi_k \subset \Pi$ such that if $\pi \epsilon \Pi' \subset \Pi$ and Π' is cardinally measurable, then $\Pi' \subseteq \Pi_k$ (and similarly for ϕ).

We may refer to any Π_k or ϕ_k as an *individual output* or *individual input*. Finally, we should add the transparent property

B10 : If $\pi, \pi' \epsilon \Pi$, then $(k\pi) \perp (k\pi') = k(\pi \perp \pi')$, and similarly for ϕ.[11]

On the basis of the above propositions we can arrive at a simple (and very general) concept of production function. Let ψ_i be the set of all concrete processes (π, φ) such that π belongs to a given individual output, Π_i. By writing $q =$ meas (π), we can describe ψ_i by a *numerical* function

$$q = q_i(\varphi). \qquad \dots \quad (10)$$

---

[8] Of course, this does not imply that there are no processes, (π_1, φ), (π, φ_1), where $\pi_1 \neq \pi$, $\varphi_1 \neq \varphi$.

[9] The last two propositions exclude creation and annihilation.

[10] They satisfy Al by virtue of B4-B6.

[11] For k rational this property follows from B4, B9 and Theorem 9.

Almost needless to say, $q_i(\varphi)$ is not a single-valued function : the same φ may be used according to various "techniques" and, above all, distributed differently among the units of production.

B11 : For a given φ, $q_i(\varphi)$ has a maximum, $Q_i(\varphi)$.

This proposition does not follow from B1-B10;[12] it is introduced only for avoiding irrelevant technicalities. A single-valued production function

$$q = Q_i(\varphi) \qquad \ldots \quad (11)$$

is thus obtained. But even this form is not very helpful, either for our own problem or for most others. The individual outputs and individual inputs do not correspond to products and factors as these terms are understood both in theory and everyday business. The present Π, for instance, does not contain an individual output representing "mutton" alone. And although every concrete input includes a management unit, management by itself can produce nothing; hence, no individual input consists solely of management.

In some cases, the distinction of an individual product or factor can be based upon the fact that a *quantitative* difference exists between two *qualitatively* different outputs or inputs. If $\pi' \epsilon \Pi_i$, $\pi'' \epsilon \Pi_j$, $i \neq j$, then $\pi' \perp \pi'' = \pi \epsilon \Pi$, i.e., π'' is the quantitative difference between π and π'. The relevant case, however, is that where the difference does not belong to Π : if $\pi =$ "1 lb wool and 3 lb mutton" and $\pi' =$ "1 lb wool and 2 lb mutton", the difference $\xi =$ "1 lb mutton" is not a concrete output. Unfortunately, nothing warrants that by taking all such *external* differences we shall ultimately derive every individual factor and individual product. Complementarity may prevent us from arriving at an external difference consisting of management alone. It is thus seen that it is not possible to derive the concept of individual factor (or product) from that of measurable inputs (or outputs). Qualitative distinctions cannot be derived from measure, for—as we have said in the introduction—quality is prior to quantity. Therefore, the concepts of individual factor and individual product must be introduced as primary concepts in any analysis of production. As we shall presently see, this point has not been always realized.

B12 : There exist two classes of cardinally measurable sets, $\{P_i\}$, $i \epsilon T$, and $\{F_j\}$, $j \epsilon H$, such that if $\pi \epsilon \Pi$ or if $\varphi \epsilon \phi$, then

$$\pi = P_{i_1} \perp p_{i_2} \perp \cdots \perp p_{i_n}, \quad \varphi = f_{j_1} \perp f_{j_2} \perp \cdots \perp f_{j_m}, \qquad \ldots \quad (12)$$

where $p_{i_k} \epsilon P_{i_k}$ and $f_{j_k} \epsilon F_{j_k}$. P_i and F_j are called individual product and individual factor respectively.

A concrete output (input) can then be represented as a vector in a space where each coordinate corresponds to an individual product (factor) and conversely. The fact that such a space may conceivably have an infinity of coordinates raises no difficulty. In a concise form, we shall write $V(f_j)$ for the vector representing a given concrete input, where f_j, $j \epsilon H$, represent the coordinates of the individual-factor space. The production functions (10) and (11) become

---

[12] As the instructive example in the author's paper "Leontief's System in the Light of Recent Results," *Review of Economics and Statistics*, 1950, p. 214, shows.

$$q = q_i(V) \quad \text{and} \quad q = Q_i(V), \qquad \ldots \quad (13)$$

respectively. They are relations between individual *factors* and a given individual *output* : therefore they are of no use in analyzing joint supply. But for our own purpose they suffice.

4. *The tautological thesis.* By this term we shall understand the thesis that "constant returns to scale is...purely a matter of definition,"[13] or in other words, that it is always possible to define the factors of production in such a way as to arrive at a production function represented by a homogeneous function of the first degree.[14] Ordinarily the argument goes no further than this assertion. Hahn, however, stands out in this respect for having tried to show how constant returns follow *tautologically* from a particular definition of factors. Consequently, his argument deserves to be analyzed in detail.

Following an old hint,[15] Hahn proposes to "classify all those units between which the marginal rate of substitution is unity into the [same] category F_i." Using our notations we can state this rule as follows: if

$$Q_i(a+V) = Q_i(b+v), \qquad \ldots \quad (14)$$

where V represents the factors kept constant, then a, $b \epsilon F_i$. (The only way in which I can interpret Hahn's "arbitrary units 'a' " is that they are some concrete factors in our sense.) Hahn goes on to say that (14) must be satisfied "independent of the size of the input, where size is measured as a fraction or multiple of the standard unit 'a'." He fails to say anything about measure, but the last statement clearly implies that all concrete factors are cardinally measurable *before* we come to apply his equivalence rule. In conclusion, we can restate Hahn's rule as follows : if $a \epsilon F_i$, $b \epsilon F_j$,

and
$$Q_i[(ka)+V] = Q_i[(kb)+V] \qquad \ldots \quad (15)$$

for any non-negative k, then *as far as economic conduct is concerned*, $F_i = F_j$.[16] That is all, according to Hahn. However, this rule obviously fails to describe the real economic conduct adequately. To begin with, (15) may be satisfied only for a particular V. If this is the case, no producer could possibly behave as if $F_i = F_j$. We must add that (15) should be satisfied not only for any k, but also for any V; this is tantamount to saying "for any size of the output." But even this amendment is insufficient : (15) may be satisfied only for some but not all products. Thus, one automobile may very well be substitutable for ten dog sleighs in Lapland: most definitely, this does not apply to Egypt. We must then add also that (15) should be satisfied for all outputs Π_i. It is clear, however, that even in this amended form the rule does not constitute a definition of individual factor. It merely imposes the condition that any distinction between the F_i's that would not matter in production should be eliminated. If, say, a certain type of worker can *always* be substituted by a certain machine without a change in output, the worker and the machine should not be listed as separate factors. But before the rule can be applied a list of F_i's, all cardinally measurable, must exist (B12).

---

[13] Hahn, *op. cit.*, p.137.

[14] This argument is endorsed also by McLeod, *op. cit.*, p. 130.

[15] See, for instance, N. Kaldor, "Mrs. Robinson's 'Economics of Imperfect Competition,' " *Economica*, 1934, p. 359.

[16] Clearly, this rule implies that F_i and F_j are perfectly divisible. But this restriction is not necessary. We need, however, the condition that if (ka) exists, then (kb) exists and vice versa, in other words, that F_i and F_j have the same cardinal scale.

The fact that Hahn believes that his definition tautologically leads to constant returns is due to a logical error in his proof. In passing from the factor combination (a, b) to (ka, kb) he made an incorrect use of the concept of rate of substitution by ignoring that this concept requires output to remain constant.[17] In fact, Hahn prefaces his proof with the remark that his definition does not preclude variable returns "to successive amounts ka of the category F_i." Strangely enough, it did not strike him that $(ka, 0)$ represents such successive amounts *and* proportional variations of factors as well.

Other arguments for the tautological thesis seem to run on a different line : any production function can be represented by a homogeneous function of the first degree if the ordinal scales of factors are appropriately chosen. Of course, the inverse transformation of (T), Example 1, changes the production function $q = e^{u^2+v^2-3}$ into $q = \sqrt{xy}$. However, this type of argument contains a rather unsuspected flaw, namely, the supposition that it is *always* possible to choose the scales as claimed. The following simple example exposes the fallacy.

Example 9 : Let $q = uv+u+v$ be the expression of a certain production function when the factors are measured on some arbitrarily chosen scales, u, v. The problem is to find the monotonic transformations $u = g(x)$, $v = h(y)$, such that $g(0) = 0$, $h(0) = 0$, and

$$q = \lambda(gh+g+h) = g(\lambda x)h(\lambda y)+g(\lambda x)+h(\lambda y), \qquad \dots \quad (16)$$

for all non-negative values of λ, u, v. Setting $y = 0$, we obtain the functional equation

$$\lambda g(x) = g(\lambda x), \qquad \dots \quad (17)$$

which has the well-known solution $g(x) = \alpha x$, α being an undetermined parameter. Similarly, $h(y) = \beta y$. If these solutions are introduced in (16), the latter yields $\alpha\beta = 0$. This shows that there are no ordinal scales for the factors such that the production function in question should be represented by a homogeneous function of the first degree.

In this connection we may recall Joan Robinson's tempting idea of transforming the natural units of all factors into efficiency units.[18] Obviously, this amounts to replacing cardinal measures by some ordinal scales chosen so as to render the production function homogeneous. The above example shows why such a proposal could but be ill-fated. This she hurried to admit at the first opportunity,[19] but it is not certain whether she realized the real drawback of her proposal: that it assumes *all* productions to be susceptible of being represented by a homogeneous function of the first degree.

---

[17] Hahn, *op. cit.*, p. 132. See also *infra*, fn 22.

[18] *Economics of Imperfect Competition*, p. 109 *passim*.

To exclude the case of production functions not reducible to a homogeneous function is not a simple matter of definition, but one concerning the structure of natural laws. For after all, any production function is subject to natural laws. By a special choice of definitions and, especially, scales we may—as indeed we often are—be able to simplify the formulae by which those laws are expressed, but it is baseless to attribute to these laws a structure of our own choice without any regard for factual considerations. This, in essence, is the position from which Chamberlin's argument proceeds. His argument may be further strengthened by the following observations: The relevant shape-properties of the production function—such as constant returns, decreasing marginal rate of substitution, etc.—are characterized by the fact that they remain invariant under one single type of scale transformation and only under that type: $x = \mu u$, $y = \nu v$. This characteristic is a consequence of the fact that in the actual world products and factors are cardinally measurable (for which reason we may refer to the mentioned properties as *cardinal* properties). Now, it stands to reason that, once the factors are listed, there is no longer any room for manipulating their definitions for the purpose of arriving at a pre-selected shape of the production function.

5. *Divisibility and homogeneity.* The striking feature of all arguments advanced by the advocates of the tautological thesis is their insistence upon the perfect divisibility of factors. All sorts of fables—such as Hahn's world of ant-men and ant-machines—are designed in order to explain how from this assumption it follows tautologically that the optimum scale of production is indeterminate. But, as Chamberlin first pointed out, perfect divisibility of factors is rather a smoke screen which obscures the issue: the optimum scale may be indeterminate even though not all factors are perfectly divisible. All we need is that all scallops in Chamberlin's diagram should be tangent to the same horizontal straight line.[20] In fact, we may go even further and, for the sake of the logical probing, assume that all factors have a natural unit. In this case, the production function $q = f(x, y)$ exists only for integral values of x and y. But that does not preclude $f(\lambda x, \lambda y) = \lambda f(x, y)$ for all integral values of λ, x, y. Neither indeterminateness nor homogeneity (of a function) are concepts restricted to the case of continuous variables. Why then the insistence upon the perfect divisibility of factors?

The answer to this question is that the usual type of argument against the existence of an optimum scale confuses divisibility of *factors* with divisibility of *processes*. To wit, Hahn writes: "since perfect divisibility [of factors] is assumed, it is solely a matter of *subdividing any single productive process* into a large number of stages [i.e., identical processes save for size], and this by definition is possible."[21] This clearly shows that Hahn's assumption of perfect divisibility applies not only to factors but also to processes.

---

[19] Joan Robinson, "Euler's Theorem and the Problem of Distribution," *Economic Journal*, 1934, p. 402n.

[20] Chamberlin, *op. cit.*, p. 243.

This tacit connection between a property of factors and one of processes implies a peculiar view of natural phenomena, to which the respective authors may or may not subscribe consciously. The view boils down to this : In nature all transformations consist of mere reshuffling of some primary substances; such reshuffling can bring about no qualities other than those inherent in the primary substances. *All processes then are indifferent to size, just as the process of measure is.* Indeed, as we have repeatedly emphasized, measuring involves some operations, i.e., it represents a process. And this process stands out as one that does not alter the quality of the element involved— its input, as it were. This is what we usually mean by "measure is indifferent to quality." The rationale of the view mentioned above seems to be that because some elements are cardinally measurable— that is, they can go through a certain process without any qualitative residual—the same must be true of all processes in which they are inputs.

Now, from the assumption that all processes can be divided or compounded without any qualitative residual it follows that all natural laws, not only economic processes, are expressed by homogeneous functions of the first degree. This *is* a tautology. It is then this assumption that is crucial for the tautological thesis. All other points upon which past controversies have spent much effort, are irrelevant. This is especially true for the assumption of "perfect divisibility of factors." Of course, the assumption is void of any meaning if factors are not cardinally measurable.

6. *Natural laws and the qualitative residual.* Examples of processes that are not indifferent to size are so abundant in natural sciences that one can only wonder how their existence may ever be ignored by other disciplines. At the microscale, organic chemistry offers innumerable examples where *new* qualities emerge after polymerization, i.e., after a certain scale is reached. For a topical illustration one may also mention the critical mass of atomic explosion. At the macroscale, in the theory of structures it is almost impossible to find a linear relation between *homogeneous and perfectly divisible* materials—iron, cement, insulation, etc—and variables expressing measures of some quality—resistance to strain, elasticity, radiation, and so on. The qualitative residual—of which we have spoken earlier —is reflected in the non-linearity of laws such as these, which relate to two distinct categories of variables: one essentially quantitative, the other pertaining to quantified qualities.

No doubt, the heat radiation from two identical spheres is twice as great as that from one such sphere. And so is the quantity of material from which they are constructed. That much everyone knows about $1+1=2$. But the issue of the optimum scale is only partly related with this *external* addition. The main problem concerns what happens when the *same* process is expanded or constructed. For simple but convincing illustrations we need not search the intricate world of life phenomena. To construct a sphere of radius, R, and wall thickness, d, it takes a quantity of materials proportional to $3R^2d-3Rd^2+d^3$; the heat radiated through such a sphere is proportional to R^2, while the heat that could be stored inside it is proportional to $(R-d)^3$. There is no linear relation between these expressions, and yet they all represent homogeneous variables. This example indirectly shows that the larger the furnace the less coal (proportionally) we need to keep it at a constant temperature. However, as the size of the sphere grows larger, the same wall thickness will no longer resist internal strain, and the sphere will collapse under its own weight. To prevent this we have to make the walls thicker and thicker; ultimately, any further increase in size becomes definitely uneconomical.

---

[21] *Op. cit.*, p. 134. Italics added.

The illustration recalls Herbert Spencer's splendid analysis of the optimum size of a bird. As he explained, volume grows faster than the area enclosing it, and the latter, faster than its average diameter. To store energy efficiently the bird's size must be large. But a large bird is heavy and beyond a certain weight the wing bones would have to be so long that they will break under the strain of lifting the body into the air. All individual processes whether in biology or technology follow exactly the same pattern : beyond a certain scale some collapse, others explode, or melt, or freeze. In a word, they ceased to work at all. Below another scale, they do not even exit.

The advocates of the indeterminateness of optimum scale, may, if they so wish, ignore all this evidence of the existence of qualitative residual in all laws. But then they must be prepared to accept another tautological conclusion, which seems to have thus far escaped notice. For if there is no qualitative residual, *there is nothing to prevent the process by which one factor is substituted for another from being indifferent to size.* Consequently, all isoquants should be linear and all production functions should be linear homogeneous expressions of factor quantities. The formula $q = \Sigma A_i f_i$ would thus represent the universal law of production.[22]

7. *Individual processes.* The fact that the authors supporting the tautological thesis mean "perfect divisibility of processes" when they say "perfect divisibility of factors", explains why they insist that every term should be defined while they fail to define "individual" process. For naturally, if processes are perfectly divisible, then *process loses all individuality* and there is nothing to define. As there is no natural unit connected with a perfectly divisible substance, there is no individual process involved in the concept of concrete process. But as we have explained in the preceding section, the issue concerning the optimum scale centers upon what happens within the *same* process. And we can speak of the same factor or the same process only in reference to an individual factor or individual process.

The difficulties connected with the concept of individual process are well-known. In spite of this, no theoretical description of reality can dispense with the primary notion of individual process : all natural laws describe in the first place an individual process. The *isolated* system of physics is nothing but such a process.

---

[22] Some points in Hahn's proof discussed earlier seem to suggest that he might have unwittingly reasoned from this position. For if the production surface is a plane the concept of rate of substitution can easily be extended to movements in any direction.

In biology and in Social sciences the concept may raise some worthy issues because at times it has been applied to such troublesome categories as species, social classes or political parties. In production theory, however, we see no reason why "an ice-cream plant" should be a less legitimate notion than "tree" or "protein molecule". There is only one point that must be clearly understood. What we regard as "one" or as "many" in each situation depends upon the problem at hand. In production theory we are led to distinguish two kinds of individual processes. The first is the process for which the term plant has been imported from biology : it is an *individual physical process* which may be assimilated to a spontaneous natural process for it is governed almost exclusively by natural laws. The second kind of individual process reflects *the unity of organization* that supervises and controls production. In addition to the natural laws governing the physical processes included in it, this latter process is subject to the laws concerning the human ability to control natural phenomena for a definite purpose.

In logical order it is the plant process that comes first; it also is the more critical of the two. A most convenient and transparent way of describing it is $q = F(W)$, the relation between output and the "material" input. This relation may vary from one process to another, even if the individual output remains the same, but broadly speaking it is in each case determined by the natural laws involved.

The following propositions express some very obvious features of the situations to which we ordinarily apply the concepts just described.

C1 : There exists a class of individual factors, $\{E_i\}$, $i\epsilon M$, $M \subset H$, such that every E_i has a natural unit, m_i, and for every concrete process (π, V), the vector $V(f_j)$ has at least one coordinate $f_j \neq 0$, $j\epsilon M$.

A unit, m_i, obviously represents a management unit. Such a unit may consist of a single person or of a board together with its general staff. There is no restriction in this respect.

C2 : Given any concrete process, (π, V), there exist concrete processes $\alpha(\pi_i, V_i)$, with $\pi, \pi_i \epsilon \Pi_k$, and

$$q = q_1 + q_2 + \ldots + q_n, \quad V = V_1 + V_2 + \ldots + V_n, \qquad \ldots \ (18)$$

and such that for every $V_i(f_j)$ the following conditions hold

$$f_a = m_a, \ a\epsilon M \quad \text{and} \quad f_j = 0 \quad \text{for } j \neq a, \ j\epsilon M. \qquad \ldots \ (19)$$

It is obvious that a process $\alpha(\pi, V)$ satisfying (19) represents a process under one management. For such a process we may put $V = (W, m_a)$, where W is a vector in the space of non-management factors, and write $\alpha(\pi, W;\ m_a)$ instead of $\alpha(\pi, V)$.

The description of a process by $\alpha(\pi, W;\ m_a)$ is however incomplete, for it includes no information regarding the number of individual physical processes covered by the firm. We must add the information that

$$\pi = \pi_1 \perp \pi_2 \perp \cdots \perp \pi_u, \quad W = W_1 + W_2 + \ldots + W_u, \qquad \ldots \ (20)$$

where $\pi, \pi_i \in \Pi_k$ and each pair (q_i, W_i) represents a concrete instance of an individual physical process. In case $u = 1$, we shall simply write $\omega(\pi, W; m_a)$ instead of $\alpha(\pi, W; m_a)$ and (20).

Definition 14 : Given Π_k, m_a, and u, the set $I_k^u(m_a)$ of all processes $\alpha(\pi, W; m_a)$ satisfying (20) is a u-plant firm. (If $u = 1$, we shall omit the upper-script.)

C3 : The set $I_k^u(m_a)$ is never empty.

C4 : If $\alpha(\pi, W; m_a) \in I_k^u(m_a)$, then for the same distribution (20) and for any $b \in M$, there exists a π' such that $\alpha(\pi', W; m_b) \in I_k^u(m_b)$.

The meaning of these two propositions is obvious: C3 states that any product can be produced by a u-plant firm whatever u and m_a may be; C4 states that any management unit can manage any process.

C5 : If $\omega(\pi_1, W_1; m_a)$, $\omega(\pi_2, W_2; m_a) \in I_k(m_a)$ there exists a π such that $\omega(\pi, W_1 + W_2; m_a) \in I_k(m_a)$.

We shall write

$$\omega(\pi_1, W_1; m_a) \oplus \omega(\pi_2, W_2; m_a) = \omega(\pi, W_1 + W_2; m_a)$$

and refer to $\oplus$ as *internal* addition.

Definition 15 : The process $\omega^+(\pi^+, W; m_a) \in I_k(m_a)$ is said to be efficient if $q^+ \geqslant q$ for any $\omega(\pi, W; m_a) \in I_k(m_a)$. The process $\omega^*(\pi^*, W; m_a) \in I_k(m_a)$ is completely efficient if $q^* \geqslant q$ for any $\omega(\pi, W; m_b) \in I_k(m_b)$, $b \in M$.

From B11 it follows that every efficient process is contained in $I_k(m_a)$. The only reason for a process $\omega(\pi, W; m_a)$ not being efficient is that the respective management unit is *intentionally* either careless in choosing the best techniques according to its knowledge or negligent in running the process thus chosen. Since elimination of this whimsical aspect of the human nature does not affect the conclusions of our argument, we may assume that all processes are efficient. The difference in efficiency between two processes then is due only to the qualitative difference of their management units, which covers knowledge of techniques as well as other management abilities.

Let us observe that if all processes are efficient the set $I_k(m_a)$ can be described by a singled-value function

$$q = U(W; m_a). \qquad \ldots \quad (21)$$

This is not identical to the usual concept of plant production function. In the first place, in $U(W; m_a)$ the unit of management is clearly set out as a qualitative unit. Secondly, in (21) the output is cardinally measurable; the formula, therefore, represents only a "slice" in the usual production function of a plant producing a multiproduct. We should also add that since we shall have frequent occasions to consider only variations of W such that $W = wW_0$, W_0 being given, it is convenient to refer to w as the scale of the process.

The question of whether a completely efficient process always exists as a concrete process is more involved and will be taken up in the next section.

8. *Proportionality, returns and efficiency.* One type of argument against the existence of an optimum scale makes use of the proposition that "if every factor is homogeneous, then doubling all factors doubles the output." We shall refer to such an argument as "the proportionality thesis." It goes without saying that there can be no quarrel about the validity of the above proposition if applied to an *external* addition, as that of B4 or C2. Controversy arises only if the doubled output is obtained by an individual producer, i.e., by a firm. In examining the possible interpretations of the proposition in question it is necessary to treat the one-plant firm separately from the multi-plant firm, for the two situations do not lead to the same conclusions. But, first, we must consider the various conceivable views on the nature and the role of the management factor. These are :

(α) Management is a fictitious factor of production introduced in economic analysis by a faulty conception of the productive process. In a static process, as in any other such process, every factor has one main role, that of keeping the process going. By definition, each factor must perform its specific role if it is to be a factor at all. Consequently, a productive process does not need a special factor to control and coordinate it. If we came to talk about a management unit it is only because of the need to identify processes; a management unit then is only a name. In this case, only W is a true variable in (21), and the latter becomes

$$q^* = U(W). \qquad \ldots \quad (22)$$

The economic process of a plant is then reduced to a purely physical process governed by technological (i.e., natural) laws.

(β) Management, even though a proper factor with a distinct role, cannot influence the relation between input and output except in a negative way by being intentionally remiss—like any other human factor. If this eventuality is excluded—as we have done— then the relation between input and output is determined exclusively by technological considerations. This leads again to (22), with the difference that m_a must now count as a cost element. We ought to add that the view just described is most unrealistic, for in a sense it attributes the quality of *perfection* to all efficient management units. We have nevertheless listed it for the sake of completeness.

(γ) Output depends upon the quality of the management unit, in the sense that each management unit can approach with various degrees of approximation the *ideal* output given by a formula such as (22). This ideal production function is determined only by technological laws and represents the individual physical process upon which the corresponding plant production rests.

In other words, q^* represents the upperbound of $U(W; m_a)$ for $a \epsilon M$. If this upperbound is reached by $U(W; m_s)$ then m_s is completely efficient for the scale W (Definition 15). However, in order to take account, first of difference between ideal laws and actual processes, and second, of the infinite variability of human quality, it is advisable to assume that no management unit is completely efficient and that there exist infinite many management units, all different in some way from each other. In particular, given W and ε there exist infinite many management units, m_a, such that $q^* - \varepsilon < U(W; m_a) < q^*$.

We can now pass to the analysis of the two cases mentioned earlier, (1) one-plant firm and (2) multi-plant firm, for each alternative α, β, γ. (1α) In this case "doubling all factors doubles the output" can mean only that $nU(W) = U(nW)$. This leads to

$$q^* = \lambda W, \qquad \dots \quad (23)$$

where λ is some constant. Nothing need be added here to what has already been said in Section 6 about the non-linearity of individual physical processes in order to see that (1α) is of no avail for the proportionality thesis. (1β) This alternative, too, leads to (23) and hence is subject to the same objections as (1α).

(1γ) In this case, "doubling all factors doubles the output" may mean that

$$nU(W; m_a) = U(nW; m_b), \qquad \dots \quad (24)$$

where n has an appropriate value and m_b is "the n-fold of m_a."[23]

The question that arises immediately is that of the operational definition of "n-fold." There seems to be no other way to go about it than the *implicit* definition :

Definition 16 : If (24) is satisfied, m_b is the n-fold of m_a and we write $m_b = n \otimes m_a$

As explained in the introductory section such a definition is useless by itself. We need an existence proposition

Z1 : Given a concrete process $\omega(\pi, W; m_a)$ and any appropriate n, there exists $b \epsilon M$ such that $\omega(n\pi, nW; m_b)$ is a concrete process.

In plain words this states that any number of one-plant firms can be replaced by a single one-plant firm without any loss of efficiency as far as the relation between the non-management input and the output is concerned. The proposition, which is far from transparent, is not of purely academic interest : it concerns the controversy over the merits of large scale production (not to be confused with large corporation.)[24]

---

[23] There is absolutely no point in considering the case of a factor which would have a cardinal scale like that of Example 8 : factors either are perfectly divisible or have a natural unit. The same is then true of the cardinally measurable set $\{nW\}$. Accordingly, in what follows we shall assume that $U(nW; m_b)$ exists either for all values of n, or for only such values that $nW = pW_0$, where p is an integer and W_0, the natural unit of $\{nW\}$.

We can easily proved that if Z1 is true, then (23) is true. Let $U(W_0)/W_0 = \lambda$ and $W = nW_0$, W being chosen arbitrarily. Let $\varepsilon > 0$ and $\varepsilon/n > \eta > 0$. Let $a \varepsilon M$ be such that $U(W_0; m_a) > U(W_0) - \eta$. By Z1, there exists $b \varepsilon M$ such that $nU(W_0; m_a) = U(W; m_b)$. Hence $U(W; m_b) > \lambda W - \varepsilon$. On the other hand, it is immediate that $U(W; m_b)$ is not greater than λW. Hence, (23) is proved. Therefore, (1γ) is to be rejected on the same ground as (1α) and (1β).

Before passing to the case of a multi-plant firm, let us point out a very important conclusion of the preceding argument. What refutes the proportionality thesis in the case of a one-plant firm is the fact that no individual physical process is indifferent to size. The point that something similar may hold for management organizations, if added, would make the argument all the stronger; but it is not needed. This is no longer true for a multi-plant firm. In a multi-plant firm the individual physical processes are added *externally* : if to a given W there corresponds the output q^* in a single individual process, an output of nq^* can be obtained from nW in n individual processes. The ideal output of a multi-plant firm can then be proportional to W (at least for multiples $W = nW_0$ of the optimum scale, W_0, of the individual process).

In the analysis of the multi-plant firm we shall distinguish the same alternatives (α)–(γ).

(2α) Since management is a pure fiction, a multi-plant firm, too, is just an identification label corresponding to our own subjective desire of referring in some way to n individual processes. It makes no difference whether some processes are thus grouped under one label or under several labels. The multi-plant firm becomes synonymous to "industry." Alternative (α) constitutes, therefore, a triumph for the proportionality thesis, even though in a definitely vacuous manner.

(2β) As any management unit is assumed *perfect*, any m_a can run any scale with *complete efficiency*. Therefore, if $\omega(\pi, W; m_a)$ is a concrete process, then $\omega(n\pi, nW; m_a)$, too, is a concrete process. By observing that from the last proposition, $m_a = n \otimes m_a$ (by Definition 16), we reach the conclusion that "doubling all factors doubles the output" is vindicated. However, let us point out that the optimum scale of the firm is, in this case, infinite, not indeterminate. Naturally, perfect management units—if they actually existed—should be exploited to the fullest extent.

(2γ) In this alternative, the proportionality thesis can be formulated as follows:

Z2. *If* $\omega(\pi, W; m_a)$ *is a concrete process, then given any* $n > 1$ *there exists* $b \varepsilon M$ *such that* $\alpha(n\pi, nW; m_b)$ *is a concrete process of an n-plant firm. By implicit definition,* $m_b = n \otimes m_a$.

---

[24] We should point out a few further issues related to Definition 16. As it stands this definition allows for the troublesome result that $m_a = n \otimes m_b$ and $m_b = n \otimes m_a$ at the same time, even if the individual out-put and the ratios between non-management factors do not change. To avoid this, we could make the definition stronger by requiring that (24) should be fulfilled for all $W = mW_0$. But this is not needed for our immediate purpose.

Nothing we have said thus far militates against this proposition. Clearly, it allows for the existence of optimum scale in the case of one-plant firm. Nor can it be rejected on the ground that in an individual process decreasing returns must sooner or later set in if one factor is invariable and all others grow proportionately. This proposition (the so-called law of variable proportions) may be invoked to prove that a one-plant firm has necessarily an optimum scale if the management unit does not change. But in Z2, the management unit is free to change with the scale of operations. This allows for the possibility that in Chamberlin's diagram one scallop should follow another to infinity, all tangent to the same horizontal line and *each representing an individual producer operating a different number of plants.*[25]

On the other hand, there seems to be no common-sense objection against the idea that the merger of any number of firms producing the same individual output should not affect the ratio between output and non-management factors, *as long as no restriction is imposed upon the size or the quality of the new management unit.*

This does not mean that Z2 vindicates the proportionality thesis. At most it enables us to state a perfect tautology or write down an elegant formula. But it does not solve the problem of optimum scale. This is a problem of efficiency, and—as we have repeatedly emphasized —efficiency measure requires cardinal measurability. For this, $m_b = n \otimes m_a$ does not suffice: we must know how many m_a's make an m_b *in the cardinal sense.* And this is an insoluble problem in *real terms.* Indeed, not all physical elements of an m_a—presidents and commissars, consultants and advisors, generals and marshals, etc, etc—are included in every other m_b. In real terms, management units differ qualitatively, and that is the end of it. For a meaningful comparison of all management units we have no other choice than to turn to prices, provided however that in the respective economic system the latter reflect adequately the relative cost to the community. For this reason Chamberlin's idea of analyzing the problem of optimum scale with the aid of a diagram of a family of average cost curves[26] seems for more promising than using the production function and its isoquants—however more respectable the latter approach may be.

CONCLUSIONS

An analysis such as ours, using formal logic blended with elementary common sense, could not claim to settle the intricate issue of the optimum size of an economic organization, which has for long been a main target in the controversies over the merits of large corporations and of large governments as well. We can only claim to have isolated the numerous related issues.

In the first place, we have shown that the ordinary concept of efficiency (as well as other equally important concepts of production theory) has no meaning if factors and products are not cardinally measurable. The issue of cardinal measurability, therefore, is far from idle. Cardinal measurability is not merely one convenient ordinal scale, but a real, specific property of numerous material variables. This property can be completely described by a series of operational propositions which are not common to all types of measurables.

---

[25] This point is extremely important in view of the fallacy—heard now and then—that the existence of the optimum size for a firm follows from the law of variable proportions.

[26] Chamberlin, *op. cit.*, p. 233.

Secondly, we have shown that the optimum size of an economic organization is not a technical problem like that of the optimum size of a plant. The optimum size of organization involves comparisons of quality; efficiency, therefore, belongs to the qualitative residual of the analysis of a productive process exclusively by the production function.

As a mere speculation, one may however regard efficiency as a quality of "thinking" matter. In this case, one may invoke the fact that relations between quantified qualities and true quantities are never linear and conclude that any organization has an optimum size, just as any plant has one. Clearly, the optimum size in both cases must change with the progress of our knowledge of "inert" and "thinking" matter.

CAPITAL FORMATION AND ITS FINANCING IN THE PRIVATE CORPORATE SECTOR IN INDIA*

By N. S. R. SASTRY and V. K. SASTRY

Bombay, India

1. INTRODUCTION

The private corporate sector in India comprising all joint stock companies has been playing an invaluable part in the development of industries. The formation of joint stock companies goes back to the middle of the nineteenth century and since then the sector witnessed a continuous growth which accelerated substantially during the first two plan periods (1951-1961); the total paid-up capital of all non-Government companies together which was about Rs. 99 crores in 1917-18 increased nearly eight-fold to Rs. 775 crores in 1950-51; by the end of the first plan period in 1956 the paid-up capital amounted to Rs. 958 crores and it expanded further to Rs. 1,200 crores by the end of the second plan period. In the pre-war period the growth of this sector was greatly facilitated by the managing agency system. The managing agents, mostly organised as firms, provided not only promotional and managerial services but also much needed finance particularly in the early days of industrialisation when neither enterprise nor capital were plentiful. Subsequently, however, the system evoked much public criticism and in the post-independence period, restrictive and controlling measures were introduced in the Indian Companies Act (1951 and 1956). Since Independence, several measures were initiated by the Government to accelerate industrial tempo of the country which helped, directly or indirectly, promote the growth of the corporate sector. An idea of the importance of the private corporate sector can be had from the fact that this sector accounted for about one-fifth of the total capital formation in the private sector during the first two plans. The object of this note is to review the trends in capital formation and its financing in the private corporate sector engaged in mining and manufacturing during the first two plan periods.

2. COVERAGE AND SOURCES OF DATA

The private corporate sector comprises mainly non-Government public limited companies and private limited companies and the basic data on the operations of these companies are provided by the annual studies on company finances made by the Reserve Bank of India and published from time to time in their monthly Bulletin. The Reserve Bank's studies on company finances are based on a purposively selected Sample of medium and large public and private limited companies, each with a paid-up capital of Rs. 5 lakhs and over; the number of public limited companies thus selected being 750 for the period 1951 to 1955 and 1,001 (which includes almost all the 750 for the earlier period) for the period 1956 to 1960 and coverage in terms of paid-up capital of the selected companies being fairly high at about 66% and 75%, respectively, for the two periods mentioned above. The selected companies cover not only mining and manufacturing industries but also such activities like trading, transport etc., and omitting the latter, the number of mining and manufacturing companies covered by the Reserve Bank's analysis is 524 for the period 1951–55 and 725 for the period 1956-60 and the present study is confined only to the operations of these companies. In terms of paid-up capital, the coverage of these companies among all non-Government public limited companies engaged in mining and manufacturing activities is about 65.7% and 81.6%, respectively, for the two periods. The private limited companies are omitted from the purview of the present study as data relating to these companies are available only for the second plan period. Further, as the number of public limited companies on which this study is based is different for the two plan periods, there may be an element of non-comparability in the data for the two periods even though such comparison is confined to certain important rates of growth and ratios but this may not be a serious limitation in view of the fairly high coverage (in terms of paid-up capital) of the selected companies.

---

* Any views expressed here are the personal views of the authors and are not necessarily those of the institution which they serve.

Rate of growth of assets. Table 1 presents the various rates of growth of assets formation as revealed by the selected companies during the period 1951–60.

TABLE 1. RATES OF GROWTH OF ASSETS FORMATION OF PRIVATE CORPORATIONS IN MINING AND MANUFACTURING INDUSTRIES

period	total gross assets	capital formation		fixed assets		inventories
		gross	net	gross	net	
1951	11.2	12.0	13.1	6.1	2.6	23.5
1952	2.2	2.4	−0.3	7.7	7.2	−6.5
1953	4.1	3.9	1.8	8.2	8.3	−4.4
1954	6.7	6.6	5.6	8.2	8.1	3.0
1955	10.0	9.6	11.1	10.4	14.0	7.7
1951-55 : average	7.8	7.8	6.9	9.6	9.4	4.5
1956	14.6	17.0	20.7	14.8	19.7	21.9
1957	12.2	15.2	17.3	17.4	22.8	10.6
1958	7.6	8.4	7.6	11.9	12.9	0.6
1959	6.4	6.2	4.5	7.5	5.5	3.1
1960	11.2	11.0	10.4	8.0	5.0	18.7
1956-60 : average	12.7	14.4	15.1	15.0	16.8	13.2

Note : Data relate to 524 companies for 1951–1955 and to 725 companies for 1956–1960 in this and all the following tables.

In the year 1951 which marked the beginning of the first five year plan there was a substantial increase in the total gross assets as well as total capital formation, bulk of which was accounted for by a very heavy speculative inventory accumulation which occurred during the year as a result of the Korean War boom. The rate of growth of gross fixed assets formation during that year was, however, much lower at 6.1 per cent. The rate of investment in gross fixed assets witnessed a sizable growth from 1952 onwards and in the last year of the first plan period it was much higher at 10.4 per cent. It may be observed in this connection that though the first plan period commenced from 1951, the actual plan proposals and efforts to implement them came out much later. The assets formation which gained momentum towards the end of the plan period continued till the middle of the second plan and again tapered off towards the end of the second plan period. The lack of a uniform trend in the growth of assets is perhaps not an unusual phenomenon in the sense that investment tends to be irregular, completion of substantial investment projects being followed generally by a slackness and it may therefore be more appropriate to compare the average rate of growth during certain periods. The average rate of growth of total gross assets (the total of physical assets and financial claims) for the second plan period is higher at 12.7 per cent per annum as compared to the corresponding rate of 7.8 per cent during the first plan period. The difference in the average rates of growth of gross fixed assets in the two plan periods is more marked; it moved up from 9.6 per cent per annum during 1951–55 to 15.0 per cent per annum during 1956–60.

In order to examine the differences in the patterns of growth among different groups of industries, we have divided all the industries covered by the present study into three groups, viz., capital goods, intermediary goods and consumer goods. Iron and steel and engineering industries are included under capital goods while mineral oil, jute textiles, cement, electricity, paper and coal are included under intermediary goods. Vegetable oils, sugar, cotton, silk and woollen textiles and matches are included under consumer goods. The average rate of capital formation of these three groups of industries is shown in Table 2.

TABLE 2. RATES OF GROWTH OF CAPITAL FORMATION : INDUSTRY GROUPS

	total gross assets		gross capital formation		gross fixed assets		inventories	
	plan period							
	first	second	first	second	first	second	first	second
1. all industries	7.8	12.7	7.8	14.4	9.6	15.0	4.5	13.2
2. capital goods	14.3	22.4	13.0	27.8	12.5	31.1	14.1	21.0
3. intermediary goods	6.6	11.8	6.8	11.8	8.2	11.3	2.1	13.8
4. consumer goods	6.0	8.8	6.6	10.5	9.4	10.9	2.9	9.8

During both the plan periods the rates of growth of capital formation as well as gross fixed assets formation were much higher in respect of capital goods industries as compared to the corresponding rates of growth of intermediary and consumer goods. The average rate of growth of gross fixed assets of capital goods industries was much higher during the second plan (31.1 per cent) as compared to the first (12.5 per cent); in the case of intermediary and consumer goods industries the rate of inventory accumulation showed a marked increase in the second plan period whereas the rate of expenditure on fixed assets registered only a moderate rise. The bulk of the investment under capital goods was under iron and steel, investment projects in which were almost completed by the middle of the second plan period.

In the context of such varying rates of growth of stock of capital in mining and manufacturing industries it would be of interest to examine the roles of different methods (internal and external) of financing the capital formation and to examine the factors that influence the choice among the alternative sources of financing. An obvious limitation of the present study is that the period is too short to permit any firm inferences about the long-term trends in the pattern of financing capital formation and the empirical findings presented below should therefore be regarded only as tentative.

Internal and external sources of finance. Internal finance represents funds obtained from the retentions out of current income. For the present study 'gross internal sources' are defined as the sum of (i) net profits retained, (ii) provision for depreciation and (iii) increase in tax reserves net of increase in income-tax advance; the last item may be regarded as roughly equivalent to tax provision during an year *less* tax paid (irrespective of whether it is by way of advance payment or otherwise) during that year. 'Net internal sources' are equal to 'gross internal sources' *less* provision for depreciation. External sources refer to funds obtained through new capital issues, bank loans, mortgages, other borrowings, trade credit, etc. The sum of internal and external sources of funds represents the total sources of funds obtained during the year and is by definition equal to total uses of funds for acquiring physical as well as financial assets.

The percentage shares of gross internal sources and external sources in the total sources of funds during the decade 1951-60 are presented in Table 3.

TABLE 3. INTERNAL AND EXTERNAL SOURCES AS PERCENTAGE OF THE TOTAL SOURCES

year	total	internal	external
1951-1955 : average	100	53.3	46.7
1951	100	51.3	48.7
1952	100	96.0	4.0
1953	100	69.6	30.4
1954	100	37.6	62.4
1955	100	51.2	48.8
1956-1960 : average	100	46.5	53.5
1956	100	37.3	62.7
1957	100	24.0	76.0
1958	100	62.0	38.0
1959	100	57.3	42.7
1960	100	54.9	45.1

The share of internal sources in the total sources of new financing shows no uniform trend during the period 1951–60. However, the average share of internal sources during the first plan period was 53.3 per cent while it was lower at 46.5 per cent during the second plan period. Profits as well as retentions were in general higher during the second plan period but the average rate of capital formation during this plan period was substantially higher as compared to the first plan period and consequently there has been a greater dependence on the external sources. It may be useful to examine what proportion of total capital formation was accounted for by the internal sources even though it may not be quite realistic to associate specific sources of funds with specific uses. The gross internal sources represent the saving or surplus on current account of the firms and the difference between capital formation and the saving represent net financial deficit of the firms, assuming capital transfers to be negligible. The figures of internal sources and of financial deficit thus arrived at expressed as percentages of total capital formation together with the corresponding rates of capital formation (i.e. fixed assets plus inventories) are shown in Table 4.

TABLE 4. PERCENTAGE SHARE OF INTERNAL SOURCES IN GROSS CAPITAL FORMATION

year	gross capital formation	internal sources	net financial deficit (+) or surplus (−)
1951-1955 : average	100	64.6	35.4
1951	100	57.8	42.2
1952	100	105.7	−5.7
1953	100	86.9	13.1
1954	100	48.4	51.6
1955	100	64.5	35.5
1956-1960 : average	100	56.0	44.0
1956	100	41.9	58.1
1957	100	24.1	75.9
1958	100	97.5	2.5
1959	100	72.6	27.4
1960	100	66.9	33.1

The average share of internal sources in total capital formation declined from 64.6 per cent during the first plan period to 56.0 per cent during the second plan period while the average rates of capital formation increased from 7.8 per cent in the first plan to 14.4 per cent in the latter period. Even though there have been substantial year to year variations in the percentage share of internal sources to total capital formation, there appears to be a general negative association between the two and the coefficient of rank correlation between these two variables works out to be significant at −0.794. In other words the years, when the rate of capital formation was higher, tended to be characterised by lower share of internal sources. It may be of interest to compare the magnitude of financial deficit (i.e. the excess of investment over saving) in relation to the investment of these Indian corporations with the corresponding figures for other countries and the relevant figures are given in Table 5.

TABLE 5. FINANCIAL DEFICIT AS PERCENTAGE OF GROSS CAPITAL FORMATION

countries	financial deficit of the corporate sector as percentage of gross capital formation
India	
1951-55	35
1956-60	44
Japan (1954-57)*	46
West Germany (1955-57)*	37
U.S.A. (1954-56)*	23

**Source :* A study on money flows in Japan 1958, The Bank of Japan.

The dependence on external sources for capital formation by Indian corporate sector during the second plan period was of the same order as in Japan but much higher as compared to West Germany and the U.S.A. As those figures seem to suggest, the extent of financial deficit in relation to capital formation appears to depend on factors other than the degree of development of an economy.

The share of internal sources in total sources of funds in respect of the three groups of industries mentioned earlier are presented in Table 6.

TABLE 6. SHARE OF INTERNAL SOURCES IN TOTAL SOURCES : INDUSTRY GROUPS

industry group	internal sources	
	plan period	
	first	second
1. all industries	53.3	46.5
2. capital goods	45.0	38.4
3. intermediary goods	62.3	48.8
4. consumer goods	62.6	47.6

All the three groups of industries registered a decline in the average share of internal sources in the total sources of finance in the second plan period as compared to the first. In both the plan periods capital goods industries depended to a larger extent on external sources than intermediary and consumer goods industries. The structure of financing as revealed by the share of internal sources was almost similar for intermediary goods and consumer goods industries. The financial structure of these three groups of industries also appears to suggest that higher the rate of growth of capital stock the higher will be the dependence on external sources.

Components of internal financing. The trends in the three components of internal sources, viz., (i) depreciation provision, (ii) net accrual of tax reserves and (iii) net retained profits may now be examined. The percentage shares of these three components during the period 1951–60 are given in Table 7.

TABLE 7. PERCENTAGE SHARE OF COMPONENTS OF INTERNAL SOURCES

year	total	depreciation provision	tax reserves	net retained profits
1951-1955 : average	100	61.9	5.2	32.8
1951	100	45.4	14.0	40.7
1952	100	102.6	—28.9	26.3
1953	100	79.9	—6.1	26.2
1954	100	90.8	12.2	—2.9
1955	100	39.3	12.1	48.7
1956-1960 : average	100	56.3	6.6	37.0
1956	100	38.1	11.0	50.9
1957	100	79.4	–4.3	24.9
1958	100	35.2	6.8	58.0
1959	100	65.5	2.2	32.3
1960	100	49.9	7.6	42.6

It will be observed that net accrual to tax reserves is of relatively smaller significance accounting for, on an average, 5 to 6 per cent of total internal sources during the two plan periods. Depreciation provision, on the other hand, accounted for a major portion of internal sources but its share had, however, declined from 61.9 per cent during 1951–55 to 56.3 per cent during 1956–60. Further, in the years in which the share of depreciation was low, profits were also generally found to be low. It may be noted that the figures of depreciation are those allowed by the companies by the conventional accounting methods and may not represent the actual amounts required for making up the loss incurred through capital consumption. In a study of the financial aspects of capital formation, changes in actual provision are of greater relevance than the actual capital consumption for it would throw light on the behaviouristic aspect as to how the actual provisions for depreciation vary at various rates of growth of the stock of capital. Depreciation provision expressed as percentage of expenditure on fixed assets is shown in Table 8.

TABLE 8. DEPRECIATION AS PERCENTAGE OF EXPENDITURE ON FIXED ASSETS

year	depreciation provision as percentage of gross fixed assets formation	rate of growth of gross fixed assets
1951-55 : average	49.7	9.6
1951	78.1	6.1
1952	53.7	7.7
1953	49.9	8.2
1954	51.3	8.2
1955	33.5	10.4
1956-60 : average	38.6	15.0
1956	26.5	14.8
1957	24.7	17.4
1958	35.1	11.9
1959	55.5	7.5
1960	63.4	8.0

The provision for depreciation expressed as percentage of gross fixed assets formation was much lower at 38.6 per cent during the second plan period than in the first plan period (49.7 per cent) while on the other hand average rate of expenditure on fixed assets was higher at 15.0 per cent in the second plan as compared to 9.6 per cent in the first plan period. If it can be assumed that the provision for depreciation represents approximately the actual loss through capital consumption, a rise in the ratio of depreciation to capital expenditure indicates, other things remaining the same, a net expansion of assets and a fall in the ratio represents greater replacement. In such a case the depreciation-expenditure ratio will be high in periods when capital expansion is low and conversely. The year to year variations in the ratio and the rate of expansion of fixed assets bear out, as expected, a close association in opposite directions and the rank correlation between the two for the period under consideration is as high as −0.97.

Turning to net profit retentions, it may be observed that their share in total internal sources was relatively higher in the second plan period as compared to the first; the average shares for the second and first plan periods being 37 per cent and 32.8 per cent respectively. In this context it may be worthwhile examining the marginal rate of profit retention. If a relationship of the type $R = \alpha+\beta P$, where R represents net profit retention and P represents net profits after tax, each expressed as a ratio to net worth, can be assumed then an estimate of β would provide a measure of the marginal rate of net saving or net profit retention of the corporate sector. Estimates of the above relationship obtained by the simple least squares method based on 17 industry aggregates for some selected years are given below:

year	relationship between R and P	coefficient of correlation
1951	$R = -2.201+0.668\,P$ (0.0613)	0.94
1954	$R = -2.734+0.637\,P$ (0.0457)	0.96
1957	$R = -4.022+0.949\,P$ (0.0660)	0.95
1960	$R = -3.326+0.812\,P$ (0.0504)	0.97

Figures in brackets are the standard errors of the respective coefficients.

It will be seen that the marginal rate of savings of the corporate sector is fairly high and it is substantially higher at 81–95 per cent in the latter two years (in the second plan period) than in the former two years at 64–67 per cent (in the first plan period). It was observed earlier that the dependence on external sources tends to increase with the rate of growth of capital stock and this taken in conjunction with the fact that the marginal rate of savings has also increased and is at a fairly high level in the latter years shows that the role of external financing in future expansion programmes is likely to assume greater significance.

The difference in the structure of internal finance among the three industry groups are brought out in Table 9.

TABLE 9. PRECENTAGE SHARE OF COMPONENTS OF INTERNAL SOURCES : INDUSTRY GROUPS

industry group	depreciation		taxation		net retained profits	
	first plan	second plan	first plan	second plan	first plan	second plan
1. all industries	61 9	56.3	5.2	6.6	32.8	37.0
2. capital goods	47.6	51.2	8.1	3.9	44.3	45.0
3. intermediary goods	47.5	53.5	5.8	8.7	40.0	37.8
4. consumer goods	66.8	58.5	1.6	6.5	31.6	35.0

It will be observed that the share of net retained profits in the total internal sources remained more or less the same at 44 to 45 per cent during the two plan periods in respect of capital goods industries, whereas in the case of consumer goods there has been a moderate increase in the share of net profits and a reverse trend was noticed in the case of intermediary goods. A significant feature was that during both the plan periods the share of net profits retained by capital goods industry is substantially larger than in respect of consumer goods industry. Depreciation provision on the other hand accounted for a relatively larger portion of total internal sources in the case of capital and intermediary goods industries as compared to consumer goods industries.

Structure of external financing. The percentage share of different sources of external financing to total external sources is presented in Table 10.

TABLE 10. PERCENTAGE SHARE OF THE COMPONENTS OF EXTERNAL SOURCES

year	total	fresh capital	total borrowings	bank borrowings	other sources
1951-1955 : average	100	17.4	46.9	15.7	35.7
1951	100	11.3	69.7	41.5	19.0
1952	100	480.2	—1234.9	—1617.4	854.7
1953	100	34.1	26.3	−5.7	39.6
1954	100	7.3	50 0	28.5	42.7
1955	100	20.1	48.9	12 7	31.0
1956-1960 : average	100	20.8	51.5	30 0	27.7
1956	100	15.9	59.7	40.8	24.4
1957	100	16.2	57.5	27.2	26.3
1958	100	26.0	50.4	17.1	23.5
1959	100	44.5	6.3	−3 9	49.1
1960	100	17.5	56.6	47.1	25.9

A striking feature revealed by the above table is the increase in the share of bank borrowings during the second plan period as compared to the first; bank borrowings which constituted about 16 per cent of total external sources in the first plan period nearly doubled their share during the second plan period. The share of fresh capital shows only a moderate rise from 17.4 per cent during 1951–55 to 20.8 per cent during the subsequent period 1956–60. It may be recalled that the above data pertain mostly to old and established companies and do not include new companies which came into existence subsequent to 1956, and as such the rise in the importance of fresh capital as an external source of finance may be much more for the corporate sector as a whole than what is indicated above. Other sources which comprise mainly trade dues and miscellaneous current liabilities accounted for about 36 per cent in the first plan period but their share had gone down to 28 per cent in the second plan period. The importance of this source would be much less if they are netted against similar acquisitions of financial assets and miscellaneous current assets.

Turning to the structure of external finances of the three groups of industries some interesting features are revealed as may be seen from Table 11.

TABLE 11. PERCENTAGE SHARE OF THE COMPONENTS OF EXTERNAL SOURCES: INDUSTRY GROUPS

	fresh capital		total borrowings		bank borrowings		other sources	
	first plan	second plan	first plan	second plan	first plan	second plan	first plan	second plan
1. all industries	17.4	20.8	46.9	51.5	15.7	30.0	35.7	27.7
2. capital goods	7.5	24.3	54.5	54.6	10.8	18.1	38.0	21.1
3. intermediary goods	26.0	26.9	34.9	42.8	10.8	29.8	39.1	30.3
4. consumer goods	14.5	9.0	67.7	59.4	26.6	48.9	17.8	31.6

During the first plan period the share of fresh capital was the lowest at 7.5 per cent in the case of capital goods industries followed by consumer goods at 14.5 per cent while the share for intermediary goods industries was little over one-fourth of the total external sources of finance. There has been considerable expansion in the share of fresh capital in the case of capital goods industries in the second plan period accounting for about 24 per cent, whereas for consumer goods industries it has shown sizable decline. In the case of intermediary goods the share of fresh capital remained more or less the same during both the plan periods. Borrowings from banks were relatively of a smaller significance in the case of capital goods and intermediary goods industries as compared to consumer goods industries during both the plan periods even though the share of bank borrowings has shown a uniformly marked increase in respect of all the three groups of industries. The share of total borrowings (including borrowings from banks) by capital goods industries remained more or less same at a little over half of the total external sources during both the plan periods. In the case of consumer goods industries even though the total borrowings accounted for the bulk of the external sources their importance had shown a moderate decline in the second plan period as compared to the first.

PLANNING AND ECONOMIC GROWTH

By J. KATZ-SUCHY

University of Warsaw

Even before I went to India, Professor Mahalanobis' work was familiar to me. I think that all of us, Poles who were trying to follow as closely as possible from distant Warsaw India's economic progress, were inclined to identify his name with Indian planning as such. This identification bears already a homage seldom accorded to planners.

Today, after five years stay in India, I can speak with a far better knowledge of the cause, and I must say, that the favourable prejudice which I brought with me to Delhi on Professor Mahalanobis' unique contribution to Indian planning and the multifarious activities of the Indian Statistical Institute, has been but strengthened. Personally, I owe to Professor Mahalanobis much of the insight I have been able to acquire on the Indian economic perspective. Talking to him had been a very rewarding experience during my stay in Delhi and the several short visits I had paid to Calcutta. Some of our conversations had a very practical follow-up and, to a certain extent, paved the way to some economic agreements presently existing between India and Poland.

Almost ten years elapsed since the discussion started on Mahalanobis' proposals for the Second Five-Year Plan. The world picture has changed considerably and today economic planning enjoys an almost universal recognition. More or less comprehensive plans are being drafted, voted and to a lesser extent enforced, in almost all the developing countries of Asia, Africa and Latin America. Planning is being commanded even by those who a few years ago were wasting a considerable amount of energy in an attempt to discredit in the eyes of the developing countries the idea of planning at large and to cast a doubt on the philosophy underlying the Indian Second Five-Year Plan in particular.

Planning today has become synonymous with economic development. Thus both problems of development and planning form today a central theme of economic thought. The time has passed when the main concern of the economic theories was the problem of equilibrium and employment. Planning is generally considered as the best strategy of economic development and a short-cut path to economic prosperity. Having realised that a repetition of the classical path of the capitalist economy is nowadays impossible, they turn their attention towards the practical experiences of the socialist policy of development characterised in many instances by slipping over certain stages considered before as necessary links in the chain of economic changes. Moreover, the experience of the socialist pattern of development and its result recounts for the growth of the conviction that economic development is impossible without planning.

Planning has become today the overall strategy of development, requiring scientists and specialists, economists, statisticians and technicians able to translate economic policy decision into practical and integrated short and long-term economic task calculated again both the needs and possibilities of the country under consideration.

This almost unanimous acceptance of the concept of planning, should not be taken at its face value. Behind semantic similarities there exists real and wide differences between, let us say, Soviet plans and the programmes hastily put together in some Latin-American countries in the spirit of the Punta del Este Charter.

Planning, however, should not be identified with state intervention or merely of programming, budgeting or earmarking of funds over a certain period for one or another purposes. Development being an integrated process acting simultaneously on many levels and on all the sectors of the national economy cannot dispose of overall planning. It must cover the basic fields of the national economy, the production of industrial goods, heavy and light industry, energy and transport, agriculture and consumption. It cannot be limited to the infra-structural field of production or those industries where the state has to assign capital because of heavy cost and slow and low return.

Planning must establish proportions between all the sectors viewed in a physical way. This includes national income, and its sources, its division between consumption and accumulation as well as its proper division between the various groups of population. It must establish proper ratios of growth, while the choice of ratios as well as a choice between accelerated expansion and less rapid growth is a policy decision and virtually has been decided by various countries depending upon their economic history, the existing levels of development and of resources, and the needs of the population who should be beneficiary of all the changes and for whom whole process of development is inaugurated. All this can be practically seen in the great battle fought over the Indian Second Five-Year Plan and its philosophy of development. The Indian Plan has been an important landmark in the worldwide approach to planning and a way to popularize planning among the countries of Asia and Africa. Not only planning as such, but also a clear understanding that the plans should be viewed in physical terms and that the solution of the ills of backwardness lies in a rapid industrialization as a base providing for the development of agriculture, for rising levels of employment and creating a dynamic internal economic market.

One could summarize Professor Mahalanobis' impact on Indian thinking on planning by stressing the following three features:

(i) the advocacy of physical planning, (ii) the grasp of the correct strategy for an accelerated industrialization with due emphasis on capital goods industries, (iii) the understanding of the necessity to proceed with some institutional changes —though, I must say it frankly, this has been the less encouraging aspect of actual Indian planning.

The body of ideas, so briefly described above, no doubt tries to incorporate many of the experiences of the socialist countries with respect to priorities of development and the type of planning. It differs from our own philosophy of development only on one but a very significant point, namely that of the scope and ways of carrying the institutional reforms.

The approach to institutional changes is the basic criterion to establish the social content of the process of development. It is that, what distinguishes the socialist pattern of development from the capitalist one and from that pattern adopted by the countries of Asia, Africa or Latin America emancipated from colonial rule.

The socialist pattern of development starts with a victorious social revolution. In Poland it coincided with a national revolution—featured by the two basic acts: the creation of the socialist sector of the economy through nationalization of industry, trade, transport and banking and an agrarian reform which abolished feudal relations in agriculture by dividing land and creating new forms of agricultural development. These two acts not only change the social content of the national economy, but set the country on a new road of development by providing the basis for the accumulation of resources for productive planned investment as well as for social expenditure for the population.

Although institutional changes are undetachable from any process of development, their character, size and form vary in various countries which entered the road of development not by socialist revolution, like India, where the emphasis on the state and public investment and the state sector becomes the most dynamic factor in economic development. Nationalisation is being applied only to part of the industry (e.g. foreign capital, mining). Agrarian reforms are, of no consequence and aim not at a land division among landless or smaller holder but like the Prussian pattern aim at creating a strong group of large farms run on the capitalistic basis.

The process of decolonisation and coming into independent being of many new countries, anxious to supplement political independence with economic development to satisfy the legitimate expectation of their peoples for higher standard of living and a fair share in the amenities of modern life has given rise to many theories and studies of the problem of economic development and made the theory of growth the most lively branch of modern economics. Many new important studies have been made which increase our knowledge of the problem and are helpful to the economist and their practitioner. But in the abundance of the theories of growth lies also their danger. Many are too schematic and many under the guise of economic theories raise political expectation to direct the developing countries towards such patterns which would permit the survival of the colonial pattern of economy in new form.

In many cases the process of growth is reduced to arithmetical exercises where given impulses lead to certain outputs. This overemphasis of simple quantitative relations leads to a serious underestimation of the institutional aspects of the process of growth. Thus stages are set according to quantitative measure of capital formation, without due attention to the institutional steps and social changes which could lead to an increase of capital formation. Such a pattern of stages of growth if applied to a developing country does not reveal anything, neither of its economic nor of its social conditions, neither of its possibilities nor of its prospects of development. To set a pattern of development of a developing country one has to review its tradition and history, the political and social forces active in the country, its resources and institutions. Stages are set not only by levels of production or the rate of capital formation, but first of all by relations of production, levels of productivity and the pattern of formation and distribution of national income. But I hasten to emphasize it, the Mahalanobis approach is certainly free from the illusion, nourished by some fashionable but oversimplified theories of growth, that the so-called "take-off" is just the matter of crossing quantitative threshold in capital formation. Economic and social progress requires qualitative changes in the institutional sphere or to put it in Marxian terms, within the realm of relations of production.

In the essay presented to Prime Minister Nehru on the occasion of his seventieth birthday, Professor Mahalanobis thought it proper to call his approach as "the middle way."* The merits and demerits of this path will be better judged by future historians, as well as the degree of correspondence between theory and actual practice. A learned discussion has already started in academic circles on the merits and demerits of Professor Mahalanobis' model. Whatever their final evaluation—both of historians and economists—may turn to be, Professor Mahalanobis' writings and advices, his research and day-to-day work at the Indian Statistical Institute and at the Planning Commission are factors which induced India to embrace and pursue the way of planned development and, the example of India, helped many other countries of Asia and Africa in choosing a similar path.

Development, planning and non-alignment are closely linked. This is stressed by India's case. Non-alignment does not mean neutrality—a situation which no developing country can afford. Non-alignment means an active participation in the solution of contemporary problem in a world full of dangers. War would mean not only a bar to the plans of development but also might endanger the very independence of the young and developing countries struggling with many difficulties. This would happen irrespective of the place of conflict. Therefore co-existence is the answer. It is the only alternative to war but also an opening of a situation where by means of international cooperation necessary economic and political changes, national liberation, economic and social progress can be achieved.

On the recognition of those principles, a fruitful policy of collaboration developed between our two countries in different spheres of activities. We work closely together in our common strivings to achieve disarmament. Economic cooperation between Poland and India is based on agreements, which in certain fields like Poland's cooperation in developing Indian coal mines extend practically up to 1980. We have given on both sides, the first modest step towards the implementation of an idea, that the present writer shares with Professor Mahalanobis, that of achieving some measure of international planning by matching national plans through long-term agreements and contracts. Cultural and scientific contacts are higher than ever, and a great post of the merit in posteriory exchanges between economist of both countries goes to the Indian Statistical Institute. Let me end this modest token of admiration presented to Professor Mahalanobis on the occasion of his seventieth birthday, by stressing once more his invaluable personal contribution to the cementing of Indo-Polish friendship.

---

* P. C. Mahalanobis (1959): Heralding a New Epoch, *A Study of Nehru*, edited by R. Zakaria p. 317.

MODELS FOR DEMAND PROJECTIONS

By RICHARD STONE

University of Cambridge

1. INTRODUCTION

Prasanta Chandra Mahalanobis has enlivened so many branches of research in the course of his scientific life that a contributor to this volume has a large number of possibilities to choose from. The topic I have selected spans two of these branches: national economic planning and the analysis of consumers' behaviour. Consumption is the principal purpose of all economic activity, and it is largely to diminish the gap between technical limits and 'natural' achievements in respect of consumption and its rate of growth that national economic planning is gaining ground all over the world. In drawing up a plan, therefore, a necessary ingredient is a budget of future consumption and we do not need to be fanatical champions of consumers' sovereignty to agree that in drawing up this budget we should take account of probable demands in the future situation.

The models I shall now describe are based on experience with British data to which they are being applied in Stone and Brown (1964), Stone, Brown and Rowe (1963) with some success. Thanks to the initiative and perseverance of Mahalanobis, the Indian National Sample Survey must by now have accumulated sufficient information to make an application to India practicable. Moreover, it should enable this to be done in an extended form which, by treating different income groups separately, would make it possible to discuss the quantitative effects of future redistributions of income.

In the following section I discuss the approach to demand studies which seems to me most relevant to problems of national planning, listing the factors we must take into account if our models are to help the planner and emphasizing the need for a comprehensive picture of future demand, broad to start with but capable of elaboration as knowledge and experience grows. In Section 3 I describe a series of interrelated models which provide the basis for the analysis of consumption in our present work on British economic growth (Cambridge, Department of Economics, 1962), in Section 4 I take up the question of statistical estimation and in Section 5 I give some results from our first attempt to use one of the models for demand projections. The paper ends with a brief conclusion and a list of works cited.

2. STRATEGY AND TACTICS

By the strategy of demand analysis I mean the influences on demand that we ought to consider and the form that these influences should take. By the tactics of demand analysis I mean the course we ought to follow in passing from an analysis of broad classes such as food or clothing to successively finer analyses of their components. Let us begin by looking at these two kinds of problem.

(1) *Strategy.* We should, I think, try to take account of the following sets of influences.

(a) *Income and prices.* These are the determining variables of classical demand theory. In introducing income we should remember that there is such a thing as saturation, and that at the moment we shall be much nearer it for some goods than we are for others. In Britain, for example, the average consumer is nearer to saturation for bread and cereals than he is for meat and nearer for television sets than he is for motorcars. In introducing prices we should remember that a change in one price, other things being equal, affects in principle all demands and not simply the demand for the good whose price has changed.

(b) *Tastes.* This kind of variable, so important in the real world, is assumed constant in classical demand theory and for that reason the theory forms an insufficient basis for empirical demand analysis. It is not merely that new goods are continually being introduced with a consequent question of their acceptance by consumers. The demand for old goods may change radically: in Britain the taste for alcohol has declined and for tobacco risen over the last hundred years.

(c) *Adaptation.* Classical demand theory implies that adaptations to a change in circumstances are instantaneous. In fact they are not, particularly in the case of durable goods the acquisition of which frequently faces the consumer with a problem of financing. This question may not be of the first importance in making long-term projections, especially in a country where durable goods form a small item in the consumers' budget. It assumes importance either where durable goods form a large item or where we are interested not only in the position five or ten years hence but also in the path by which this position is approached.

(d) *The distribution of income.* The classical theory of demand is largely concerned with the individual consumer and has little to say on the problem of aggregation. In practice it is usual to include in demand analysis a single income variable, the mean of the income distribution. In our work on British economic growth we have not up to now gone further than this, but it would be possible to treat households in different ranges of income separately. This would make it possible to discuss the quantitative effects of income redistribution and this may be of particular interest in the context of Indian planning.

(2) *Tactics.* Our ultimate aim is to analyse demand in as much detail as possible. But even if we had data on the fine structure of demand we should almost certainly find it difficult to obtain a consistent picture of changing demand patterns

if we were to analyse commodities one at a time. I suggest, therefore, that we should approach our ultimate aim in stages : first, analysing the main categories of demand such as food, clothing, housing, transport etc.; second, analysing the main components of these categories such as cereals, meat, fats, vegetables and so on in the case of food; and, finally, analysing the elements of the main components such as bread, flour, biscuits, cakes, cereal products and so on in the case of cereals. By this means we may hope to keep the whole picture in focus and ensure that at each stage our estimates are consistent, a point which is important for planning purposes where components must add to totals.

3. SOME COMPUTABLE MODELS

In economics, as in other subjects, we may look to models for insight or for usable, quantitative results. To be usable, a model must satisfy three criteria : it must provide an acceptable approximation to the process we are studying; it must be capable of application with existing data or data which could be collected for the purpose; and it must be computable, that is, it must be of such a form that its parameters can be estimated by a reasonably objective and efficient method. Economists, who all too often concern themselves exclusively with insight, are apt to relax if their model passes the first test; for the user of their results the second and third are equally indispensable.

With this in mind I shall now develop a particular model of demand which I have elsewhere (Stone, 1954) called the linear expenditure system. Despite its simplicity it can easily be modified to give a reasonable approximation to the pattern of demand and therefore provides a good point of departure.

(1) *The simple linear expenditure system.* In matrix notation, the equation of this system can be written in the form

$$\hat{p}e = b\mu + Bp. \qquad \dots \ (1)$$

In (1) e denotes a vector of quantities demanded, p denotes a vector of the corresponding commodity prices, and $\hat{p}$ denotes a diagonal matrix formed from p so that $\hat{p}e$ denotes a vector of expenditures. The symbol $\mu \equiv p'e$ (in which the prime indicates transposition) denotes total expenditure and b and B denote respectively a vector and a matrix of parameters. If total expenditure is divided between ν commodity groups, then the vectors have ν elements and the matrix is of order ν. Equation (1) expresses the theory that each element of expenditure is a homogeneous linear function of total expenditure and of each of the elements of the price vector.

The number of elements in b and B together are $\nu(\nu+1)$. However, these cannot all be independent. If we premultiply (1) by the unit row vector, $i' \equiv [1, 1, \dots, 1]$, we should obtain μ on the right-hand side. Thus $i'b = 1$ and $i'B = [0, 0, \dots, 0]$. This reduces the number of independent constants to be estimated to (ν^2-1).

This is still a large number : if $\nu = 10$ there are 99 constants to be estimated. Moreover, if we tried to estimate them with the amount of information usually available we could be certain that the results would be highly implausible. The way to overcome this difficulty is to introduce further restrictions on the parameters and the obvious one is to require that the matrix of elasticities of substitution be symmetric. As applied to the individual consumer, this is the Slutsky condition. It is not strictly applicable to the average consumer because of differences in individual preferences and possible changes in the distribution of income; but I shall ignore this objection. In this case, as shown in Stone (1954), B must satisfy the relationship

$$B\hat{p}(I-ib') = (I-bi')\hat{p}B'. \qquad \ldots \quad (2)$$

If B is to be independent of p, this means that

$$B = (I-bi')\hat{c} \qquad \ldots \quad (3)$$

where c denotes a vector of constants.

So, with this simplification, the number of independent constants to be estimated is reduced to $2\nu-1$: with $\nu = 10$ this means 19 constants in place of the 99 mentioned above. Naturally, this simplification is bought at a price; how high this price is, we shall see in a moment.

By combining (1) and (3) we can write

$$\begin{aligned}\hat{p}e &= b\mu+(I-bi')\hat{c}\,p \\ &= \hat{p}c+b(\mu-p'c). \qquad \ldots \quad (4)\end{aligned}$$

We can now see clearly the theory of consumers' behaviour which this system of equations represents and we can interpret the parameters b and c. According to (4) consumers buy quantities, the elements of c, at the current prices and the total cost of this expenditure is $p'c$. The money they have left over, $\mu-p'c$, they devote to the different commodities in proportion to the elements of b. Thus the elements of c can be interpreted as purchases to which the average consumer is committed and the elements of b can be interpreted as the proportions in which he devotes his uncommitted expenditure to the different commodities.

It has been shown elsewhere (Stone, 1954, 1961; Stone and Croft-Murray, 1959) that this model has a number of properties, some useful, some restrictive. The more important of these are as follows.

(a) The model is associated (Geary, 1949-50) with an ordinal utility function of the form

$$\upsilon = \varphi[\prod_{\xi=1}^{\nu}(e_{\xi}-c_{\xi})b_{\xi}] \qquad \ldots \quad (5)$$

where υ denotes total utility and φ denotes an arbitrary monotonic function. Thus υ is a function of a geometric average of the excesses of actual purchases over the corresponding committed quantities.

(b) The model is associated (Klein and Rubin, 1947-48) with a constant-utility index-number of the cost of living μ_1^*/μ_0, where μ_1^* takes the form

$$\mu_1^* = p_1'c+(\mu_0-p_0'c)\pi_b. \qquad \ldots \quad (6)$$

In (6) π_b is a geometric average of the price ratios (period 1 divided by period 0) with the elements of b as weights. Thus the amount of money the average consumer must have in period 1 to enable him to obtain the same utility as in period 0 is equal to the cost of the committed quantities valued at the prices of period 1 plus the uncommitted expenditure of period 0 changed by a geometric index of the price ratios.

(c) The model can handle only a set of competing commodity groups; it cannot allow for complementary or inferior groups (Stone, 1954).

(d) The Engel curves connecting e with μ for fixed values of the other variables are linear with intercept $c_\zeta - b_\zeta(p'c)/p_\zeta$ and slope b_ζ/p_ζ.

(e) The ordinary price-quantity demand curves connecting e_ζ with p_ζ for fixed values of the other variables are hyperbolae. With positive c_ζ they cannot be elastic.

(2) *A price-sensitive variant.* The above model takes account only of total expenditure and the price situation. Experience shows that committed expenditure is a large part of the total so that the quantities demanded are not very sensitive to prices. For large commodity groups this is plausible but clearly it would not be satisfactory in handling a set of close substitutes such as different brands of a given product or the demands for a given product from different sources of supply. In both these cases it seems reasonable to assume a fairly stable commitment to purchase a given amount of the product in total but one would expect the allegiance to a particular brand or source to be sensitive to relative prices.

If we return to (2) we can see that this kind of situation can be accommodated in the model. The symmetry of this equation would not be affected if we replaced $\hat{c}$ in (3) by $\hat{c}+\hat{p}D$ where $D = D'$. In this case we should replace (4) by

$$\begin{aligned} \hat{p}e &= b\mu+(I-bi')(\hat{c}+\hat{p}D)p \\ &= \hat{p}(c+Dp)+b[\mu-(c'+p'D)p]. \qquad \ldots \quad (7) \end{aligned}$$

The addition of D without any further restrictions than the condition of symmetry would add $\nu(\nu+1)/2$ parameters to the system, making in all $(\nu^2+5\nu-2)/2$ or, with $\nu = 10$, 74. Some simplification is, therefore, desirable, and one might approach the problem as follows.

Let us consider D as made up of three components : (i) a constant scalar, λ say, which represents the importance of price differences in effecting substitutions; (ii) a scalar, varying from year to year which ensures the correct dimensionality of D, say the inverse of the weighted arithmetic average of prices, $i'c/p'c$; and (iii) a set of elements which reflect the transfer potential between the brands or sources. For example, with three commodities we might write

$$D = \frac{\lambda i'c}{p'c}\begin{bmatrix} -c_1(c_2+c_3) & c_1c_2 & c_1c_3 \\ c_2c_1 & -c_2(c_1+c_3) & c_2c_3 \\ c_3c_1 & c_3c_2 & -c_3(c_1+c_2) \end{bmatrix} \qquad \ldots \quad (8)$$

where $c = \{c_1\ c_2\ c_3\}$. In this form D is not only symmetric but $Di = D'i = \{0\ 0\ 0\}$. Thus, if $p_1 = p_2 = p_3$, the term Dp in (7) disappears and we return to (4). If p_1 rises in relation to p_2 and p_3 then the first element of Dp is negative and the second and third elements are positive and together equal in magnitude but opposite in sign to the first element. Thus the effect of a rise in p_1 relative to p_2 and p_3 is to transfer part of the total existing commitment from brand or source 1 to brands or sources 2 and 3. At the same time there is a change in uncommitted expenditure, but this is likely to be small.

This variant involves only one additional parameter, giving 20 in all for $\nu = 10$. I have not experimented with this model but I think it might be useful in analysing the demand for a product which may come from many sources such as home production and imports from different regions abroad.

(3) *The introduction of changes in tastes.* Let us now return to the model set out in (4). Bearing in mind the meaning of b and c it is to be expected that these parameters will change over time due mainly to the fact that a rising standard of living will lead to higher levels of commitments and to a redistribution of uncommitted expenditure over the different commodities. This is an important aspect of changes in taste and can be represented by making b and c functions of predetermined variables. In the simplest case we can make them functions of time, in which case for year θ we have

$$b_\theta = b^* + b^{**}\theta \qquad \ldots \quad (9)$$

and

$$c_\theta = c^* + c^{**}\theta. \qquad \ldots \quad (10)$$

If b and c in (4) are replaced by the expressions in (9) and (10) we have $2(2\nu-1)$ independent constants or 38 if $\nu = 10$. The reason is that in (9) the additivity condition $p'c \equiv \mu$ requires that $i'b^* = 1$ and $i'b^{**} = 0$.

As indicated in Stone, Brown and Rowe (1963) this model works reasonably well and could be made to work better if the linear trends in (9) and (10) were replaced by more complicated trends. It is a simple matter to introduce quadratic trends and this will allow for tendencies to accelerate or decelerate. Such a modification yields a better fit but it is not ideal since an accelerating trend can hardly continue indefinitely and a decelerating trend can hardly be expected to become negative. Positively skewed sigmoid trends would probably be best from an economic point of view but I do not know how these could be fitted.

(4) *A link with budget studies.* In Stone, Brown and Rowe (1963) the model consisting of (4), (9) and (10) was used to analyse consumers' expenditure in Britain, divided into eight groups, over the period 1900 to 1960. Fourteen war years were omitted and so $8 \times 47 = 376$ observations were available to determine $2 \times 15 = 30$ independent constants.

Had regular budget studies been available over the period, an attempt could have been made to obtain from them extraneous estimates of the b_θ. We can see from

(4), (9) and (10) that the elasticity at time θ, $H_{\zeta\theta}$ say, of expenditure on commodity ζ with respect to total expenditure, is given by

$$H_{\zeta\theta} = \frac{\partial(p_{\zeta\theta}\, e_{\zeta\theta})}{\partial\mu_\theta} \cdot \frac{\mu_\theta}{p_{\zeta\theta}\, e_{\zeta\theta}}$$

$$= \frac{b_{\zeta\theta}\, \mu_\theta}{p_{\zeta\theta}\, e_{\zeta\theta}}$$

$$= \frac{\partial e_{\zeta\theta}}{\partial\mu_\theta} \cdot \frac{\mu_\theta}{e_{\zeta\theta}} . \qquad \ldots \quad (11)$$

From (11) we see that $H_{\zeta\theta}$ is equal to the ratio of the proportion of uncommitted expenditure devoted to commodity ζ at time θ divided by the proportion of total expenditure devoted to that commodity at the same time. We also see that the expenditure elasticity and the quantity elasticity are equal. If therefore we had estimates of $H_{\zeta\theta}$ we could obtain extraneous estimates of $b_{\zeta\theta}$ from the relationship

$$b_{\zeta\theta} = H_{\zeta\theta}\, \frac{p_{\zeta\theta} e_{\zeta\theta}}{\mu_\theta} . \qquad \ldots \quad (12)$$

In this way we should use the time series to determine only the 2ν parameters necessary to estimate the c_θ. Alternatively we might attempt a set of maximum likelihood estimates using time series and budgets together (Durbin, 1953).

I mention these possibilities because the National Sample Survey might provide the means of applying in India the model obtained by combining (4), (9) and (10). Without long, comparable series one could not base this model solely on annual data. But if the b_θ can be obtained from cross-section data, even ten years should provide enough information to form usable estimates of the c_θ.

That this possibility is worth serious consideration is exemplified by the following example taken from Stone, Brown and Rowe (1963). In Britain a large budget study was conducted in 1938 and, in the case of food, comparable data are also available for several postwar years. Estimates of the elasticity of food expenditure with respect to total expenditure derived from (11) are set out in the following table and compared with the independent estimates derived from budget studies.

TABLE 1. TOTAL EXPENDITURE ELASTICITIES FOR FOOD IN BRITAIN

	$b_{\zeta\theta}$	$\frac{p_{\zeta\theta}\, e_{\zeta\theta}}{\mu_\theta}$	elasticities from model	elasticities from budgets
1900	0.314	0.327	0.96	—
1938	0.166	0.292	0.57	0.59
1955	0.100	0.343	0.29	0.30
1960	0.080	0.309	0.26	0.25

Budget estimates for 1938, reached by different methods can be found in Aitchison and Brown (1957) and Stone and others (1954). Those for 1955 and 1960 are taken from U. K. Ministry of Agriculture, Fisheries and Food (1960). These results are striking, partly for the close agreement between the two sets of estimates and partly for the ability of the time-series model to show a strong downward trend in the elasticity although the proportion of total expenditure devoted to food was comparatively constant. The explanation of this phenomenon is that committed purchases of food were gradually rising through time while the proportion of uncommitted expenditure devoted to food was falling.

(5) *A disaggregated model.* So far, I have only considered models which relate to the average consumer. In Britain the available time series relate to consumers as a whole and are subdivided only by commodity. But in India, with the National Sample Survey, it is presumably possible to obtain the necessary data for many subgroups of the population : rich and poor, urban and rural, etc. In this way, a very full analysis of consumers' behaviour would be possible.

In deriving Engel curves at different times from budgets it seems best to use a sigmoid curve. For Britain we have obtained good results with the log-normal integral curve proposed by Aitchison and Brown (1957). This curve has a saturation level which is related in a simple way to the consumption and corresponding elasticity at any point on the curve. With regular information from budgets, it is possible to form a time series of saturation levels; from the above relationship it is possible to do this without budget data, as the following example shows.

TABLE 2. FOOD SATURATION IN BRITAIN
(£ 1938 per head)

year	observed food expenditure	percent of saturation	implied saturation
1900	21.3	0.39	54.7
1913	21.7	0.51	42.6
1924	23.9	0.59	40.5
1938	27.4	0.65	42.2
1950	31.5	0.78	40.7
1960	34.7	0.86	40.4

The most striking feature of this table is the comparative constancy of the saturation level from 1913 onwards. Even the small movements of this level may be economically determined: the relatively high level in 1938, for example, is associated with a relatively low level of food prices. The comparatively high saturation level in 1900 may be due to the less accurate information for the past or to the very simple trends in the parameters *b* and *c* which obviously could not be expected to continue indefinitely. On the other hand it may represent a real effect since it is not hard to imagine a fall in Britain in the notion of what constitutes a desirable amount of food, between the nineteenth and twentieth centuries.

The table is also interesting from the point of view of welfare. A community becomes better off as what it has approximates what it wants. This may happen either because what it has remains constant and what it wants declines or because what it wants remains constant and what it has rises. We see both tendencies at work in Britain in the last two generations.

The figures are also enlightening for the planner. It would be a mistake to regard these indirect estimates of saturation levels as accurate and immutable but at least they tell us something about what people want but do not have.

(6) *A model of adaptation.* Let us now return to the simple model of equation (4) and ask how it can be generalised to take account of the fact that adjustments to changing circumstances cannot always be carried out quickly but must, in some cases, be spread over several years. This version of the model is based on Stone and Rowe (1957) and has been developed in Stone (1960), Stone and Croft-Murray (1959). The treatment here is brief but, I hope, sufficient.

We start by dividing the purchases of each commodity into two parts : consumption denoted by an element of u and net investment denoted by an element of v. Thus

$$e \equiv u+v \qquad \dots \quad (13)$$

where $e_\zeta = u_\zeta$ and $v_\zeta = 0$ if ζ denotes a perishable good. Consumption is defined by a reducing-balance depreciation formula, so that

$$u \equiv \hat{n}^{-1}s+\hat{m}^{-1}e \qquad \dots \quad (14)$$

where s denotes a vector of written-down stocks at the beginning of the year, $\hat{n}^{-1}$ is a diagonal matrix of depreciation rates and $\hat{m}^{-1}$ is a diagonal matrix whose elements depend on n. If ζ denotes a perishable good, then $s_\zeta = 0$ and $n_\zeta = m_\zeta = 1$. If on the other hand ζ denotes a durable good, then $s_\zeta \geqslant 0$ and $1 < n_\zeta < m_\zeta$. Equation (14) thus defines consumption as the depreciation on the initial written-down stock plus the depreciation on the purchases of the year. If these purchases take place uniformly then it can be shown (Stone and Rowe, 1957) that

$$m_\zeta = \sum_{\theta=1}^{\infty} (1/\theta n_\zeta^\theta) \Big/ \sum_{\theta=2}^{\infty} (1/\theta n_\zeta^\theta) \qquad \dots \quad (15)$$

an expression which converges rapidly to $2n_\zeta - 1/3$ as n_ζ increases.

Adaptation is allowed for by relating net investment to the difference between the initial stock, s, and a desired stock s^*, at present unspecified. If the relationship is in all cases one of proportionality then we can write

$$v = \hat{r}(s^* - s) \qquad \dots \quad (16)$$

where r_ζ is the proportion of the discrepancy $s^*_\zeta - s_\zeta$ which is removed in a year and so can be interpreted as an adjustment rate.

In equilibrium, consumption, u^*, is proportional to the stock, s^*, and net investment, $v^* = 0$. Accordingly,

$$u^* = (1-\hat{m}^{-1})^{-1}\hat{n}^{-1}s^* = \hat{k}s^* \qquad \ldots \quad (17)$$

say. By combining (14), (16) and (17), we can see that

$$u = \hat{r}^*u^* + (I-\hat{r}^*)\hat{k}s \qquad \ldots \quad (18)$$

where $r^* \equiv \hat{n}\,\hat{m}^{-1}r$. It can be shown (Stone and Rowe, 1957) that

$$\lim_{n_\zeta \to 1} k_\zeta s_\zeta = E^{-1}e_\zeta \qquad \ldots \quad (19)$$

where $E^\theta e_\zeta(\tau) \equiv e_\zeta(\tau+\theta)$. Thus the whole system can be written as

$$u = \hat{r}^*u^* + (I-\hat{r}^*)x \qquad \ldots \quad (20)$$

and so

$$e = \hat{r}^*\hat{m}u^* + (I-\hat{r}^*\hat{m})x = \hat{r}\hat{n}u^* + (I-\hat{r}\hat{n})x \qquad \ldots \quad (21)$$

where $x_\zeta = [m_\zeta/n_\zeta(1-m_\zeta)]s_\zeta$ if ζ is durable and $x_\zeta = E^{-1}e_\zeta$ if ζ is perishable.

The final step is to replace u^* in (20) and (21) by an observable expression. Thus let

$$\hat{p}u^* = b\mu + (I-bi')\hat{c}p \qquad \ldots \quad (22)$$

so that

$$\hat{p}u = \hat{r}^*b\mu + \hat{r}^*(I-bi')\hat{c}p + (I-\hat{r}^*)\hat{p}x \qquad \ldots \quad (23)$$

from (20) and (22), and

$$\hat{p}e = \hat{}\hat{n}b\mu + \hat{r}\hat{n}(I-bi')\hat{c}p + (I-\hat{r}\hat{n})\hat{p}x \qquad \ldots \quad (24)$$

from (21) and (22).

This model enables us to distinguish between long and short-term responses. Thus from (22) and (24),

$$\frac{\partial e}{\partial \mu} = \hat{r}\hat{n}\frac{\partial u^*}{\partial \mu} \qquad \ldots \quad (25)$$

and so, if $\hat{r}\hat{n} = I$, the vector of short-term responses is equal to the vector of long-term responses and, in fact, we have returned to the model given by (4). But any component of $\hat{r}n$, say $r_\zeta\ n_\zeta$ may be $\lessgtr 1$ and so the corresponding short-term response may be $\lessgtr$ the long-term response.

Given n, the model expressed by (24) involves $3\nu-1$ parameters. As we shall see in the next section, the adding-up theorem can no longer ensure that $i'b = 1$ when the parameters are fitted by the simple method of least squares and so it is necessary to introduce an undetermined multiplier to ensure that this condition is fulfilled.

(7) *The analysis of subgroups.* In principle there is no limit to the number of groups among which total consumption is divided in any of the preceding models provided that each group is a substitute for every other. In practice, however, it may be impossible to carry out the calculations involved if the number of groups is large and so it may be necessary to analyse first the main groups and then their components. This can be done because the linear expenditure system is decomposable, as can be seen as follows.

Consider a group, j say, of the complete system. The equation for this group can be written as

$$\hat{p}_j e_j = \hat{p}_j c_j + b_j(\mu - p'c) \qquad \ldots \quad (26)$$

where the suffix j denotes that the vector to which it is attached contains elements relating only to group j. If we premultiply (26) by i' we obtain

$$p'_j e_j = \mu_j = p'_j c_j + i'b_j(\mu - p'c) \qquad \ldots \quad (27)$$

so that

$$\mu - p'c = (i'b_j)^{-1}(\mu_j - p'_j c_j). \qquad \ldots \quad (28)$$

If we substitute for $\mu - p'c$ from (28) into (26) we obtain

$$\hat{p}_j e_j = \hat{p}_j c_j + b_j(i'b_j)^{-1}(\mu_j - p'_j c_j). \qquad \ldots \quad (29)$$

From an analysis of the main groups we can obtain estimates of $i'b_j$ and $i'c_j$. From an analysis of the components of group j we can obtain estimates of c_j and $b_j\,(i'b_j)^{-1}$. We can thus estimate b and c for the whole system.

(8) *A link with the consumption function.* The linear expenditure system enables us to relate expenditure on individual categories to total expenditure. If we could relate total expenditure to disposable income, that is, income after taxes, we could relate expenditure on individual categories to disposable income. Following Stone and Rowe (1962) I shall now show how this can be done in a way which links up with recent developments of the consumption function.

Let us denote real expenditure by e, real disposable income by ρ and real wealth by ω; and let us divide each of these into a permanent and a transient component, denoted respectively by the suffixes 1 and 2. These components are not observable, but I shall make the assumption that a change in the permanent component of income is a constant proportion, λ say, of the excess of this year's observed income over last year's permanent component. I shall also make an identical assumption for wealth. Thus

$$\Delta^* \rho_1 = \lambda(\rho - E^{-1}\rho_1) \qquad \ldots \quad (30)$$

and

$$\Delta^* \omega_1 = \lambda(\omega - E^{-1}\omega_1) \qquad \ldots \quad (31)$$

where the operator E has the same meaning as before and $\Delta^* \equiv (1 - E^{-1})$.

Let us now assume that the permanent component of expenditure, ϵ_1, is a homogeneous linear function of the permanent components of income and wealth. Then

$$\epsilon_1 = \alpha_1\omega_1+\beta_1\rho_1 \qquad \dots \quad (32)$$

say, and that the transient component of expenditure, ϵ_2, is a homogeneous linear function of the transient component of income. Then

$$\epsilon_2 = \beta_2\rho_2. \qquad \dots \quad (33)$$

From (30) through (33) it follows that

$$\begin{aligned}\epsilon &= \alpha_1\omega_1+\beta_1\rho_1+\beta_2\rho_2 \\ &= \alpha_1\lambda\omega+\beta_1\lambda\rho+\beta_2(1-\lambda)\Delta^*\rho+(1-\lambda)E^{-1}\epsilon \qquad \dots \quad (34)\end{aligned}$$

In (34) ϵ is expressed in terms of observable variables and the parameters can be estimated without ambiguity.

Let us denote the consumers' price index by $\pi = \mu_1^*/\mu_0$ in the symbolism of subsection (1.b) above, and let us identify total committed expenditure $p'c$ with the permanent component of expenditure expressed in money terms, $\pi\epsilon_1$.

Then

$$\begin{aligned}p'c &= \pi\epsilon_1 = \pi(\alpha_1\omega_1+\beta_1\rho_1) \\ &= \pi\lambda[\alpha_1\sum_{\theta=0}^{\infty}(1-\lambda)^\theta E^{-\theta}\omega+\beta_1\sum_{\theta=0}^{\infty}(1-\lambda)^\theta E^{-\theta}\rho] \qquad \dots \quad (35)\end{aligned}$$

from (30) through (32). We can now rewrite (4) as

$$\hat{p}e = \hat{p}c+b[\pi(\epsilon-\epsilon_1)] = pc+b(\pi\epsilon_2). \qquad \dots \quad (36)$$

This is equivalent to saying that the expenditure on each class of commodity is a homogeneous linear function of the price index of that class and of the current value of the transient component of total expenditure. In this formulation, the elements of b and c can be estimated from a set of simple regressions on two variables. If (35) holds then these estimates satisfy the constraint $i'b = 1$.

From (36) we can express the expenditure on each class in terms of prices and of the current values of present and past incomes. Thus

$$\hat{p}e = \hat{p}c+\pi\alpha_2 b[(1-\lambda)\rho-\sum_{\theta=1}^{\infty}(1-\lambda)^\theta E^{-\theta}\rho]. \qquad \dots \quad (37)$$

I do not know how this model works out in practice beyond the fact that, for Britain, (34) seems to give fairly satisfactory results.

4. STATISTICAL ESTIMATION

The models we are using in our present work on the British economy are those given by (4), (9) and (10) on the one hand and by (24) on the other. In both cases the parameters can be estimated by an iterative two-stage least-squares procedure. In the first case the adding-up theorem ensures that the constraint $i'b = 1$ is satisfied; in the second case it does not. I will now give a proof of the adding-up theorem and then describe the two methods of estimation.

Nicholson (1957) propounded an adding-up theorem of considerable value in regression analysis. In its general form, this theorem states that if a set of variables subject to a linear constraint is made to depend linearly on one of their numbers and on any other variables whatsoever, then the least-squares regression estimator will be such that the linear constraint is satisfied by the regression estimates which will equal the true value in the case of the included variable.

This theorem can be proved as follows. Let Y, of type $\tau \times (\nu+1)$ denote a set of τ observations on $\nu+1$ variables such that

$$Ya = \{0, 0, \ldots, 0\} \qquad \ldots \quad (38)$$

where a is a vector of $\nu+1$ constants. As an example consider the regression on total expenditure of expenditures on individual groups of commodities. If the first column of Y represents total expenditure and the remaining columns represent group expenditures, then $a = \{-1, 1, \ldots, 1\}$ indicating that the group expenditures add up to the total. If the variables in Y are regressed on a set of variables in X, we can write the regression equation as

$$Y = XB+W \qquad \ldots \quad (39)$$

where B is a matrix of regression coefficients and W is a matrix of disturbances. The regression estimator, B^*, obtained by the method of least squares, is

$$B^* = (X'X)^{-1}X'Y. \qquad \ldots \quad (40)$$

Estimates for any period θ can be written as

$$Y_\theta^{*\prime} = X_\theta' B^* \qquad \ldots \quad (41)$$

where $Y_\theta^{*\prime}$ and X_θ' are row vectors whose elements are the values for period θ of the variables appearing in Y and X. Substituting for B^* from (40) into (41) and post-multiplying by a, we obtain

$$Y_\theta^{*\prime} a = X_\theta' (X'X)^{-1}X'Ya = 0 \qquad \ldots \quad (42)$$

from (38). Thus the regression estimates satisfy the same constraint as the observations. If the first column of Y, total expenditure in the example, is the same as the first column of X, then the first column of B^* is $\{1, 0, \ldots, 0\}$ and so the constraint is satisfied by the true value of total expenditure and the regression estimates of its components.

Let us now turn to the two estimation procedures, which I will call Method 1 and Method 2.

Method 1 : The problem is to estimate the parameters b^*, b^{**}, c^* and c^{**} in the model given by (4), (9) and (10). We begin by guessing values of b^* and b^{**} which I shall denote by b_0^* and b_0^{**}. We then form a vector of type $\nu \times 1$, y_θ say, as follows

$$y_\theta \equiv \hat{p}_\theta \, e_\theta - (b_0^* + \theta b_0^{**})\mu_\theta. \qquad \dots \quad (43)$$

We also form a matrix of order ν, Y_θ say, as follows

$$Y_\theta \equiv [I - (b_0^* + \theta b_0^{**}) i'] \hat{p}_\theta. \qquad \dots \quad (44)$$

Apart from a random element, y_θ and Y_θ are connected by the relationship

$$y_\theta = [Y_\theta \vdots \theta Y_\theta] \begin{bmatrix} c^* \\ \cdots \\ c^{**} \end{bmatrix}. \qquad \dots \quad (45)$$

If we now define $\quad y \equiv \{y_1, y_2, \dots, y_\tau\} \qquad \dots \quad (46)$

and $\quad Y \equiv \{Y_1, Y_2, \dots, Y_\tau\} \qquad \dots \quad (47)$

we can write, apart from a random element,

$$y = Xg \qquad \dots \quad (48)$$

where $X \equiv [Y \vdots \theta Y]$ and $g \equiv \{c^* \vdots c^{**}\}$. The least-squares estimator, g_1, of g is

$$g_1 = (X'X)^{-1} X'y. \qquad \dots \quad (49)$$

Given g, we can form a vector of type $\nu \times 1$, w_θ say, as follows

$$w_\theta \equiv \hat{p}_\theta \, [e_\theta - (c_1^* + \theta c_1^{**})] \qquad \dots \quad (50)$$

and a matrix of order ν, W_θ say, as follows

$$W_\theta \equiv [\mu_\theta - p_\theta' (c_1^* + \theta c_1^{**})] I. \qquad \dots \quad (51)$$

It will be noticed that W_θ is a scalar matrix. Apart from a random element, w_θ and W_θ are connected by the relationship

$$w_\theta = [W_\theta \vdots \theta W_\theta] \begin{bmatrix} b^* \\ \cdots \\ b^{**} \end{bmatrix} \qquad \dots \quad (52)$$

so that if we define $\quad w \equiv \{w_1, w_2, \dots, w_\tau\} \qquad \dots \quad (53)$

and $\quad W \equiv \{W_1, W_2, \dots, W_\tau\} \qquad \dots \quad (54)$

we can write, apart from a random element,

$$w = Zh \qquad \dots \quad (55)$$

where $Z \equiv [W \vdots \theta W]$ and $h \equiv \{b^* \vdots b^{**}\}$. The least-square estimator, h_1, of h is

$$h_1 = (Z'Z)^{-1} Z'w. \qquad \dots \quad (56)$$

Given h, we can return to (45), replace b_0^* and b_0^{**} by b_1^* and b_1^{**} and calculate the next approximation $g_2 \equiv \{c_2^* \vdots c_2^{**}\}$ of g. If we continue in this way until the estimates cease to change, we shall have reached a solution.

In estimating h, since W_θ is a diagonal matrix, the system of equations breaks down into ν separate equations. At the same time the adding-up theorem ensures that $i'b^* = 1$ and $i'b^{**} = 0$ for the estimated values of b^* and b^{**}. In estimating g on the other hand, since Y_θ is not a diagonal matrix, the ν equations all contribute to giving a single, average estimator of g.

This method is laborious and probably impracticable without electronic computing equipment. The example in Stone, Brown and Rowe (1963) involved eight consumption groups observed in each of forty-seven years. Twelve complete cycles were needed to obtain a reasonable degree of convergence and the number of operations involved was of the order of twelve million. Some of the results have been given above. We are now working on a variant with quadratic trends which seems likely to give even better results. The computing scheme is essentially the same as the one set out above.

Two main problems have to be faced in applying this method. First, we have assumed that b_0^* is equal to the vector of average expenditures on the different commodity groups and that $b_0^{**} = 0$. Our test of sufficient convergence is that $i'|\Delta c^*| \leqslant 10^{-4}$ and since we stop the calculations when this condition is satisfied, we cannot be certain that the process really is converging to values close to our final estimates though we may be fairly confident that the process does in fact converge (Stone, Aitchison and Brown, 1955). In two-dimensional problems it is often found that convergence requires that the initial estimates lie in a certain region. Sometimes it is possible to specify this region. I do not know of any results which would be helpful in the present case.

Second, in applying simple, rather than generalised, least squares, it is necessary to assume the usual scalar variance matrix for the disturbances. This is not very plausible since some commodity groups are much smaller than others and, as time goes on, prices rise and recent expenditure levels are substantially larger than those of the past. The results, however, do not indicate markedly bad fits for small groups or for early periods and, accordingly, we have not so far attempted a generalised treatment.

Method 2 : The problem here is to estimate the elements of r, b and c in (24), it being assumed that the elements of n are given. Since the adding-up theorem cannot help us in this case we must impose the condition that $i'b = 1$.

We begin by guessing values of r and b which I shall denote by r_0 and b_0. We then form a vector of type $\nu \times 1$, y_θ say, as follows

$$y_\theta \equiv \hat{p}_\theta[e_\theta - (I - \hat{r}_0\hat{n})x_\theta] - \hat{r}\hat{n}b_0\,\mu_\theta. \qquad \ldots \quad (57)$$

We also form a matrix of order ν, X_θ say, as follows

$$X_\theta \equiv \hat{r}_0\hat{n}(I - b_0 i')\hat{p}_\theta. \qquad \ldots \quad (58)$$

As with Method 1, let $y \equiv \{y_1, y_2, \ldots, y_\tau\}$ and let $X \equiv \{X_1, X_2, \ldots, X_\tau\}$. Then apart from a random element we can write

$$y = Xc \qquad \ldots \quad (59)$$

whence the least-squares estimator, c_1 say, of c is

$$c_1 = (X'X)^{-1}X'y. \qquad \ldots \quad (60)$$

Now let us define

$$w_\theta \equiv \hat{p}_\theta(x_\theta - c_1) \qquad \ldots \quad (61)$$

$$Z_\theta \equiv (\mu_\theta - p'_\theta c_1)I \qquad \ldots \quad (62)$$

and

$$\hat{z}_\theta \equiv \hat{n}^{-1}\hat{p}_\theta(\hat{x}_\theta - \hat{e}_\theta). \qquad \ldots \quad (63)$$

Also, let $w \equiv \{w_1, w_2, \ldots, w_\tau\}$, $Z \equiv \{Z_1, Z_2, \ldots, Z_\tau\}$ and $\hat{z} = \{\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_\tau\}$. Then apart from a random element we can write

$$w = [Z \vdots \hat{z}] \begin{bmatrix} b \\ \cdots \\ \hat{r}^{-1} i \end{bmatrix} \qquad \ldots \quad (64)$$

and this must be solved subject to $i'b = 1$. If λ denotes an undetermined multiplier, the normal equations are

$$\begin{bmatrix} Z'w \\ \cdots \\ \hat{z}w \\ \cdots \\ 1 \end{bmatrix} = \begin{bmatrix} Z'Z & Z'\hat{z} & \vdots & i \\ \hat{z}Z & \hat{z}^2 & \vdots & 0 \\ i' & 0 & & 0 \end{bmatrix} \begin{bmatrix} b \\ \cdots \\ \hat{r}^{-1}i \\ \cdots \\ \lambda \end{bmatrix} \qquad \ldots \quad (65)$$

from which the required estimates can be made. If we now replace r_0 and b_0 by r_1 and b_1 in (59) we can calculate the next approximation, c_2 say, of c. As before, we continue until the successive estimates cease to change.

This method is subject to much the same comments as Method 1 : it gives interesting results but it is laborious and converges slowly. No doubt the slow convergence is at least partly due to the general interdependence of economic phenomena coupled with the small samples which the analyst of economic time series is forced to work with. If cross-section data could be combined with time series in estimating the parameters of these models, the procedure would not only be greatly simplified but much less information would be required from the time series.

5. AN EXAMPLE OF DEMAND PROJECTIONS

The models and methods I have been describing have been developed for use in a model of economic growth which has been described in Cambridge, Department of Applied Economics (1962) and is being elaborated in further papers in the same series. I shall now describe the method of demand projection proposed and illustrate it by some provisional results for 1970, set out in greater detail in Stone, Brown and Rowe (1963).

In using the model consisting of (4), (9) and (10) for purposes of projection, five steps are involved.

(i) We calculate b_θ and c_θ for $\theta = 10$, that is for 1970.

(ii) We calculate p_θ for $\theta = 10$. Eventually these prices will emerge as shadow prices worked out in the general model, but those shown below are simply the result of extrapolating exponential trends fitted to data for 1950 through 1960.

(iii) We calculate μ^*_{10} from

$$\mu^*_{10} = p'_{10}c_{10}+(\mu_0-p'_0\,c_{10})\pi_{b_{10}}. \qquad \dots \quad (66)$$

This equation is the same as (6) with b and c put at their 1970 values. Accordingly μ^*_{10} represents the amount of money which would be needed by the average consumer of 1970, faced initially with the total expenditure and price structure of 1960, to feel equally well off under the price conditions of 1970.

(iv) We calculate $\mu_{10} = \mu^*_{10}e^{10\rho}$ where ρ denotes the instantaneous rate of growth in real consumption per head that we wish to consider between 1960 and 1970.

(v) We calculate e_{10} from (4) replacing b, c, p and μ by b_{10}, c_{10}, p_{10} and μ_{10}.

The original calculations were made with eight commodity groups including two very small ones, communications and entertainment. I have here reduced the number to six, combining communications with transport and entertainment with other. The results could be presented in many ways; here I will select three of them.

Table 3 shows the change from 1970 in the average consumption of each of the six groups of commodities for different values of ρ.

TABLE 3. PROJECTED CONSUMPTION LEVELS FOR BRITAIN IN 1970
(Observed levels in 1960=1.00)

	$\hat{e}_0^{-1}e_{10}$ for different values of ρ		
	0.020	0.030	0.040
food	1.10	1.11	1.13
clothing	1.43	1.72	2.04
household	1.25	1.40	1.56
transport and communication	1.39	1.67	1.98
drink and tobacco	1.14	1.23	1.34
other	1.34	1.54	1.76
total	1.24	1.39	1.54

We can see from this table that the projections for food and for drink and tobacco are well below the average but that food hardly rises at all with ρ whereas drink and tobacco rise considerably. At the same time the category 'household' almost exactly keeps pace with the average. Finally, the remaining three categories start above the average and rise more than it does.

Table 4 shows the composition of consumption and of expenditure in 1970 for different values of ρ compared with the composition in 1960.

TABLE 4. COMPOSITION OF CONSUMPTION AND EXPENDITURE IN BRITAIN: 1960 AND 1970
(Percentages)

	consumption				expenditure			
	1960	1970 for different values of ρ			1960	1970 for different values of ρ		
		0.020	0.030	0.040		0.020	0.030	0.040
food	29.4	26.0	23.7	21.6	30.9	31.0	28.5	26.2
clothing	9.8	11.3	12.1	12.9	10.0	9.4	10.1	10.9
household	23.4	23.6	23.6	23.7	21.0	22.1	22.4	22.6
transport and communication	12.5	14.0	15.1	16.1	10.7	13.2	14.3	15.3
drink and tobacco	9.7	8.9	8.7	8.4	13.2	9.6	9.4	9.2
other	15.1	16.3	16.9	17.3	14.2	14.7	15.3	15.8
total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

In this table, the consumption figures are expenditures valued throughout at 1938 prices. The left-hand side of the table shows quite marked shifts in the proportions, especially for high values of ρ. The percentages on the right-hand side show smaller shifts between 1960 and 1970; indeed for $\rho = 0.020$, approximately the normal peace-time growth rate over the last half century, the two sets of figures are very similar except that the positions of the transport group and the drink and tobacco group are interchanged. The total expenditure elasticities vary slightly with ρ and would also vary if the assumption about relative prices were changed. In the case of food, this elasticity, on the assumptions made, works out at about 0.15 in 1970.

These results should be regarded as little more than an illustration. The price projections are crude, nothing more in fact than extrapolations of past trends designed to make this illustration possible. The trends in b and c are linear, the simplest possible form. No attempt has yet been made to subdivide the commodity groups into components or to discuss the results obtained with producers or distributors. Nevertheless I think these first results justify the use and development of the models described in this paper and suggest that they might fruitfully be applied to other countries as well as to Britain.

6. CONCLUSIONS

In concluding this description of our work on demand projections, I should like to explain the links with other forms of demand analysis and the way in which I think it should be developed and used.

First, the models I have described are intended to give a plausible approximation to the changing pattern of demand, in which account is taken both of responses to income and prices and also of changes in tastes. A recognition that relationships do change seems to me essential in econometric work designed to help in planning our future. Indeed it is essential to all econometric work which aspires to being more than a numerical branch of economic history.

Second, the models are intended to be used in conjunction with a general model of economic growth which deals with such matters as production, accumulation and foreign trade as well as with consumption. The emphasis is, therefore, on consistency and manageability. No doubt, for certain commodities better approximations could be achieved, but the range that could be covered in this way is limited and it would be impossible to build a general model of consumption out of them. At present there is a need for broad, comprehensive models and for detailed studies of individual components. We must try to find methods of combining them.

Finally, results of the kind presented here, even if greatly elaborated and improved, should be regarded primarily as a basis for discussion with people of practical experience who may be able to bring additional knowledge to bear on conclusions which, with our best endeavours, are based on simple assumptions and general tendencies. We need not despair on this account. Each part of the practical world knows far more about its little corner than we do; but it does not possess a general picture and, by itself, has no means of reaching one. The task of the econometrician, in my opinion, is first to work out a detailed quantitative statement of future demand, supply or whatever it may be, in a form in which outside knowledge can be brought to bear on it, and then to reconstruct a coherent statement after the initial coherence has been destroyed by outside criticism.

REFERENCES

AITCHISON, J. and BROWN, J. A. C. (1957): *The Lognormal Distribution*, Cambridge University Press.

CAMBRIDGE, DEPARTMENT OF APPLIED ECONOMICS (1962): A computable model of economic growth No 1, in a *A Programme for Growth*, Chapman and Hall, London.

DURBIN, J. (1953): A note on regression when there is extraneous information about one of the coefficients. *J. Amer. Stat. Ass.*, **48**, No. 264, 799-808.

GEARY, R. C. (1949-50): A note on 'A constant-utility index of the cost of living'. *The Review of Economic Studies*, **XVIII** (1), No. 45, 65-6.

KLEIN, L. R., and RUBIN, H. (1947-48): A constant-utility index of the cost of living. *The Review of Economic Studies*, **XV**, No. 38, 84-87.

NICHOLSON, J. L. (1957): The general form of the adding-up criterion. *J. Roy. Stat. Soc.*, Series A (General), **120**, 1, 84-85.

Samuelson, Paul A. (1947-48): Some implications of "linearity". *The Review of Economic Studies*, **XV** (2), No. 38, 88-90.

Stone, Richard (1954): Linear expenditure systems and demand analysis: an application to the pattern of British demand. *The Economic Journal*, **LXIV**, No. 255, 511-27.

——— (1960): A dynamic model of demand (in Polish). *Przegląd Statystyczny*, 7, No. 3, 255-70. (English version, Department of Applied Economics reprint series, No. 167).

——— (1961): *Input-Output and National Accounts*, O.E.E.C., Paris.

Stone, Richard, Aitchison, J. and Brown, J. A. C. (1955): Some estimation problems in demand analysis. *The Incorporated Statistician*, 5, No. 4, 165-77.

Stone, Richard and Brown, Alan (1964): Behavioural and technical change in economic models. To be published by the International Economic Association.

Stone, Richard, Brown, Alan and Rowe, D. A. (1963): *Demand Analysis and Projections for Britain: 1900-1970. A Study in Method.* To be published for the Association Scientifique Européenne pour la Prévision Economique à Moyen et à Long Terme by the North Holland Publishing Co., Amsterdam.

Stone, Richard and Croft-Murray, Giovanna (1959): *Social Accounting and Economic Models*, Bowes and Bowes, London.

Stone, Richard and Rowe, D. A. (1957): The market demand for durable goods. *Econometrica*, 25, No. 3, 423-43.

——— (1962): A post-war expenditure function. *The Manchester School of Economic and Social Studies*, **XXX**, No. 2, 187-201.

Stone, Richard, and Others (1954): *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom*, 1920-1938, 1, Cambridge University Press.

U. K. Ministry of Agriculture, Fisheries and Food (1960): *Domestic Food Consumption and Expenditure*: 1960. H.M.S.O., London.

INDIA SINCE INDEPENDENCE
SOME RANDOM THOUGHTS

By S. SUBRAMANIAN
Director-General of Commercial Intelligence and Statistics,
Government of India

The last World War left undivided India in a state of economic distress comparable to that of other countries in South East Asia which were ravaged by military operations of a more active character. Means of production had been switched over to meet the requirements of the fighting forces in different corners of the world. Imports were conspicuously absent. The total supplies available for civilian consumption had thus been drastically cut. The financial machinery was wholly geared up to produce the money needed for Government transactions the volume of which was determined not by internal factors but by a global strategy determined elsewhere. There was, however, a relieving feature in the picture of shortages and deficiencies in that employment was kept at a relatively high level in all fields connected with the prosecution of the war.

2. The end of the war brought about a reduction in the size of the armed forces but it was only much later that a commensurate degree of contraction was envisaged in the civilian categories. Nevertheless, in the background of an increasing population, it could hardly be a matter for satisfaction if employment merely remained stationary without a corresponding expansion in connected fields. The population problem in India which has always remained intractable showed some deceptive signs of easing as a result of the falling birth rate but much of this fall was neutralized by a steeper decline in the death rate.

REHABILITATION

3. Most important of all, the forced displacement of population which emerged as a concomitant of Partition left a host of resettlement problems in its wake and to-date the absorption of the inward flow has proved one of the most pressing amongst them. The creation of new townships to accommodate the surplus and the provision of assistance to the needy in the shape of money grants and in kind by way of facilities for readjustment occupied the prior attention of the new State which had to surmount multifarious difficulties caused by the shortage of administrative personnel left behind by the withdrawal of the British.

4. The problem of resettlement was one of considerable importance in the Punjab and the smoothness with which the allotment of land was put through is an example of the efficacy of the sympathetic zeal of a friendly administration. Dandakaranya remains the symbol of similar resettlement and rehabilitation in respect of the population moving in across the eastern frontier of the country.

ABSORPTION OF STATES

5. No less important than the fulfilment of the obligation to rehabilitate displaced persons was the effective strategy adopted to integrate as many as 600 Principalities over the length and breadth of the country, some being as large in area as a major province, others being little more than tiny villages. Although the integration of such a motley group of territories was a political operation, the beneficial effect of their absorption in the body politic was the significant improvement in the availability and reliability of basic economic and social information. In some areas, in fact, before integration hardly anything was known of even the total area not to speak of the area under crops or yield. The administration was far below the expected norms of efficiency and effectiveness. With the unification of the country came rationalisation of the districts which are really the lynchpin of the entire set-up. An offshoot of this reform still remained in the shape of the centrally administered territories like Tripura and Manipur.

FRANCHISE

6. Another memorable land-mark was the extension of the franchise to all adults. Before Independence, the privilege of voting in elections was confined to the lucky few with educational and property qualifications. One of the earliest steps taken by the Constituent Assembly of independent India was to put into practice the principle of equality of all men and women of knowledgeable age. The results of the Census of population which were languishing before recent years were utilised ingeniously to demarcate constituencies for electing one or more representatives to the Legislature. The special position of the intellectual classes was not forgotten. Provision was made for Second Chambers in several States with membership chosen from special constituencies like graduates and local authorities. It is said that accelerators without brakes are likely to lead to disequilibrium. The Second Chamber thus provides a stabilising factor without going counter to democratic requirements. So far, there have been three general elections and the smooth and peaceful way they were conducted is a tribute not only to the peace-loving nature of the entire people but also to the unquestioned efficiency of the electoral machinery.

ECONOMIC AND SOCIAL DEVELOPMENT

7. Coming in the wake of the separation of Burma a decade earlier, the Partition of the rest of the country queered India's economy overnight from near self-sufficiency in foodgrains and in the principal industrial raw materials to a new dependence on imports. The country had also to look for now markets for millions of yards of cotton piecegoods. To cope with the developing situation, Planning become indispensable. The distinguishing feature of the policies adopted in this context by the National Government was the accent on democratic planning of the progress of the entire population aiming not only at increasing production all round but also distributing the income more equitably than before. It is true that much more emphasis has been laid in the plans on economic gains for society than on building up a solid social welfare system. The investment in agriculture, industry,

transport and communications together far outweighs the allotment to education, health and allied heads. In such matters of policy in regard to priorities, however, it is easy to be unduly critical; but with limited resources, clearly, first things must come first. A controversy developed on the relative importance of heavy and light industries, of large and small industries and the like. While the First Plan was inevitably concerned with strengthening the agricultural base, the Second Plan had to tackle the problems of co-ordination of the different sectors of industrial activity. The concept of mixed economy guaranteed a rightful place for the so-called private sector although the Industrial Policy Resolutions adumbrated by the State did give the appearance of inflexibility. The merit of the Second Plan was that a place was found for even village industry and a substantial allocation made for its development even while emphasising the importance of steel and other large industries. Looking back over the period of the First and Second Plans and the first two years of the Third Plan, it is evident that all the economic indicators have shown reasonably forward moves.

INDICATORS OF ECONOMIC AND SOCIAL DEVELOPMENT IN INDIA

(Base adjusted to 1948-49=100)

items	1950-51	1955-56	1960-61	1961-62
1. per capita income at constant prices	100	108	118	118[4]
2. agricultural production[1]	..	117	139	..
3. foodgrains production	101	132	159	..
4. value of mineral output[2]	112	147	257	275
5. industrial production[2]	97	132	184	195
6. electricity generated	112[2]	188[2]	370	425
7. imports-quantum	93	102	158	149
8. exports-quantum	123	125	118	123
9. wholesale commodity prices	109	90	122	122
10. consumer prices (working class)	103	98	126	130
11. railways—tonne-k. metres of goods	119	162	239	245
12. railways—passenger-k. metres carried	107	100	124	132
13. railways—net earnings	100	115	195	199
14. aviation passengers carried—internal[2]	130	123	219	246
15. aviation freight carried—internal[2]	668	796	649	051
16. membership of primary co-operative societies	120	136	272	..
17. working capital of co-operative societies	126	213	596	..
18. infant mortality rate per 1000 of population[3]	133	103	87	..
19. expectation of life at birth (years)[3]	32	..	45	..
20. percentage literacy[3]	17	..	24	..
21. employment in factories[2]	125	132	159	..
22. employment in railways[2]	97	107	126	126
23. employment in mines[2]	101	109	121	..

[1] 1949-50 as base.

[2] Figures are for calendar years with 1948=100.

[3] Actuals.

[4] Provisional.

Note : The sign .. indicates that figures are not available.

DATA COLLECTION

8. In the achievement of this moderate progress, the strengthening of the machinery for fact-finding has played an important part. No plan for progress is practicable unless the basic data are forthcoming; and the data cannot be collected satisfactorily unless there is a plan for the purpose. Before Independence, a decentralised system of data collection intimately related to the executive responsibilities of different Departments was in vogue. Immediately after Independence, the need for a co-ordinating agency was realised and steps were taken to unify the technical direction of surveys, experimentation and administrative compilation under the overall guidance of a Statistical Adviser to the Cabinet. Quite smoothly and silently, the State base has been strengthened so that now every State and even centrally administered area has arrangements for systematic analysis and interpretation of figures obtained through different channels. Attention is being given at the same time to the concepts underlying the collection of facts.

9. Statistical personnel were quite scarce before Independence. Provision was therefore made for the recruitment and training of adequate numbers of people at various levels of work. As compared with just a few hundred statistical employees in the whole country before Independence, nearly ten thousand people with basic qualifications and with sound training have now been posted for work right from the Central and State headquarters down to the district and even to the community block level. The out-flow of data is far more accurate and timely to-day than it was in 1948. The administration is no longer chary of figures but on the other hand it cries for more and better statistics. In thus building up an intellectual base for progress, a decisive role has been played by Professor Mahalanobis in the crowning years of his three score and ten.

PROJECT APPRAISAL : A TRADITIONAL APPROACH

By J. TINBERGEN
Netherlands Economic Institute

1. NATURE OF PROBLEM

One phase in the construction of a development plan consists of the selection of a programme of projects out of a larger collection of projects. The aim of this selection process is to find a set of projects which together will use the quantities available of scarce production factors and which will maximize the contribution to the aims of development policy. We will assume that it is known beforehand which factors are scarce. The outcome of the selection process may show, however, that one or more of these factors are not actually scarce. We will also assume that the aims of development policy have been defined, can be measured and have pre-conceived "prices." Such aims may be to make a contribution to national product (usually considered to be the most important aim), to make a contribution to employment, to change the inequalities in distribution, to add to the population's health and culture and perhaps still other. The "prices" to be attached to the units of each of these aims may, in principle, be fixed by taking interviews from policy-makers. This means that the contribution made to the joint aims of development policy by each single project i can be expressed in one figure y^i. This is not true of the total use of scarce factors by each project, since the relative prices of these factors cannot be determined autonomously as in the case of the aims. These prices, often indicated as "shadow prices" must reflect the relative scarcities of the factors and they depend on the selection of projects made. Indicating the relative prices of factors 2, 3, etc. in terms of factor 1 by p_2, p_3 etc. we may state that only after these prices have been chosen can one calculate the total value of all scarce factors used. The criterion to be used for the selection of projects may be the "yield" h^i of each project i which can be written

$$h^i = \frac{y^i}{f_1^i + p_2 f_2^i + p_3 f_3^i + \cdots} = \frac{1}{x_1^i + p_2 x_2^i + p_3 x_3^i + \cdots}$$

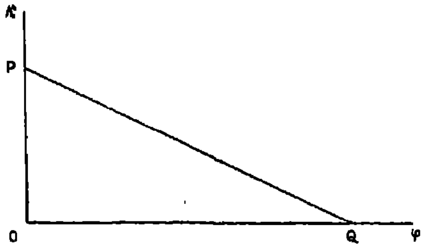
where f_1^i, f_2^i etc. are the total quantities of factors 1, 2, etc. used in the project and x_1^i, x_2^i etc. the factor coefficients, that is, the quantities of each factor used per unit of contribution to the aims. If the first scarce factor is capital and the aim of development is output then x_1 is the capital-output ratio. If the second scarce factor is foreign currency, x_2 is the foreign currency-output ratio etc. By selecting the projects in the order of decreasing h, starting with the highest value, we will get the maximum contribution out of our programme of projects; this presupposes that the p's are such as to enable us to apply exactly the quantities available of each scarce factor if we take all the projects showing an $h^i \geqslant h$. Data of our problem are the total quantities

$F_1, F_2, \ldots$ of the factors available; unknowns are the relative factor prices p_2, p_3 etc. and the limit h until which we must proceed. Formulated in this way the problem of selection is a simple problem in N unknowns, where N is the number of factors, and N equations, stating that the total quantity absorbed of each factor equals the quantity available. In this shape the problem is not a maximum problem, since no degrees of freedom exist. The maximum problem is a "built-in" maximum problem by our rule of selecting the projects in the order of their h^i.

Even so the problem has not been given yet an explicit formulation, since the total quantity of each factor used by the programme cannot always be expressed as an explicit function of the values h and $p_2, p_3 \ldots$. This is only possible if the collection of projects can be represented by an explicit frequency distribution of the projects' characteristics y^i, x_1^i, $x_2^i \ldots$. It is a modern tendency not to use such explicit frequency distributions, since they will either be very crude approximations only or have to be very complicated. The modern trend is to develop less general, numerical methods of a trial-and-error method and to aim at the exact answer. With very "irregular" frequency distributions this is probably the only possible way. In this article I propose to be old-fashioned and to use approximative general methods. For almost "regular" frequency distributions these are much easier. We lack information as to whether for practical cases the new or the old approach is more efficient. The old method has the advantage of some generality and elegance of a traditional mathematical character.

2. TWO SCARCE FACTORS: CONSTANT DENSITY DISTRIBUTION OF PROJECTS

In the tradition of successive approximations we will first tackle the problem in its simplest form. Let there be two scarce factors, capital and foreign currency, in given quantities K and V respectively. The factor coefficients will be written as $\varkappa$ and φ and our assumption is that projects are so distributed over the $\varkappa$, φ plane that the amount of y (contribution to aims) to be obtained from them is evenly spread over this plane, at least in the relevant portion of it. The quantity obtained per unit of surface, or density, is assumed constant and equal to A. By the choice of a relative price p of foreign currency in terms of capital and a limiting value of the yield h we select a set of projects falling within a triangle (cf. Figure given below) OPQ, where



Figure

the slope of PQ corresponds with p and the level of PQ with h; we have $OP = \frac{1}{h}$ and $OQ = \frac{1}{hp}$. It is an easy problem of calculus to express explicitly in h and p the total contribution Y and the quantities K and V absorbed for given h and p; we find :

$$Y = A \int_0^{1/ph} d\varphi \int_0^{1/(h-p\varphi)} d\varkappa = \frac{1}{2} \frac{A}{ph^2} \qquad \ldots \quad (2.1)$$

$$K = A \iint_\Delta \varkappa d\varphi \;\; d\varkappa = \frac{1}{6} \frac{A}{ph^3} \qquad \ldots \quad (2.2)$$

$$V = A \iint_\Delta \varphi \, d\varphi \;\; d\varkappa = \frac{1}{6} \frac{A}{p^2h^3} \qquad \ldots \quad (2.3)$$

where the latter two integrations must be carried out over the same triangle as the first one. The problem of finding p and h is extremely simple now; we find :

$$p = \frac{K}{V} \qquad \ldots \quad (2.4)$$

and

$$h^3 = \frac{AV}{6K^2}. \qquad \ldots \quad (2.5)$$

We also express Y in terms of the data and find an interesting "production function"

$$Y = \frac{3^{2/3}}{2^{1/3}} (AKV)^{1/3}. \qquad \ldots \quad (2.6)$$

3. MORE SCARCE FACTORS : CONSTANT DENSITY OF PROJECTS

The result just obtained can be generalized for any number N of factors. Using the notation introduced in Section 1 we find for the contribution Y and the quantities F_n of scarce factor number n

$$Y = \frac{A}{N!h^N} \prod_{n=2}^{N} \frac{1}{p_n} \qquad \ldots \quad (3.1)$$

and

$$F_n = \frac{A}{(N+1)!p_n h^{N+1} \prod\limits_{n'=2}^{N} p_{n'}}. \qquad \ldots \quad (3.2)$$

The solution of the problem can be written as

$$p_n = \frac{F_1}{F_n} \qquad \ldots \quad (3.3)$$

$$h^{N+1} = \frac{A \prod\limits_{n=2}^{N} F_n}{(N+1)!F_1^N} \qquad \ldots \quad (3.4)$$

and the "production function" as :

$$Y = \frac{(N+1)^{N/(N+1)}}{(N!)^{1/(N+1)}} A^{1/(n+1)} \prod_{n=1}^{N} F_n{}^{1/(N+1)} \qquad \ldots \quad (3.5)$$

of which formula (2.1) through (2.6) are special cases for $N = 2$.

If by aggregation we want to combine some of the scarce factors into one (say, qualified labour of all types) we may do so by defining the quantity used of this factor by

$$F = \prod_{1}^{m} F_n. \qquad \dots \quad (3.6)$$

The production function may then be written :

$$Y = CF^{m/(N+1)} F_{m+1}^{1/(N+1)} \dots F_N^{1/(N+1)}. \qquad \dots \quad (3.7)$$

This function approaches the Cobb-Douglas production function in two ways :

(i) the sum total of the exponents approaches 1 and

(ii) the exponents are no longer equal among them.

4. TWO SCARCE FACTORS : POWER SERIES FREQUENCY DISTRIBUTION

In an attempt to generalize our results to more complicated frequency distributions it is an established tradition to develop the frequency density function into a power series :

$$A(\varphi, \varkappa) = \sum_m \sum_n A_{mn} \varphi^m \varkappa^n. \qquad \dots \quad (4.1)$$

Our method now requires the computation of integrals

$$J(m, n) = \int_0^g \varphi^m d\varphi \int_0^{p(g-\varphi)} \varkappa^n d\varkappa = \frac{p^{n+1} g^{m+n+2}}{n+1} I(m, n+1) \qquad \dots \quad (4.2)$$

where, for simplicity's sake, we use the following symbols :

$$g = \frac{1}{ph}, \text{ hence } \frac{1}{h} = pg \qquad \dots \quad (4.3)$$

and 1

$$I(m, n) = \int_0^1 z^m (1-z)^n dz. \qquad \dots \quad (4.4)$$

It is clear that

$$I(m, n) = I(n, m). \qquad \dots \quad (4.5)$$

By partial integration we can derive a recursive formula :

$$I(m, n+1) = \frac{n+1}{m+1} I(m+1, n) \qquad \dots \quad (4.6)$$

and from (4.5), (4.6) and the value

$$I(n, 0) = \int_0^1 z^n dz = \frac{1}{n+1} \qquad \dots \quad (4.7)$$

we finally derive the general formula for

$$I(m, n) = \frac{m!\, n!}{(m+n+1)!}. \qquad \dots \quad (4.8)$$

These auxiliary results can be used to derive formulae for Y, K and V:

$$Y = \sum_m \sum_n A_{mn} \frac{m!\ n!}{(m+n+2)!} \frac{1}{h^{m+n+2}\ p^{m+1}} \qquad \ldots \quad (4.9)$$

$$K = \sum_m \sum_n A_{mn} \frac{m!\,(n+1)!}{(m+n+3)!} \frac{1}{h^{m+n+3}\ p^{m+1}} \qquad \ldots \quad (4.10)$$

$$V = \sum_m \sum_n A_{mn} \frac{(m+1)!\,n!}{(m+n+3)!} \frac{1}{h^{m+n+3}\ p^{m+2}}. \qquad \ldots \quad (4.11)$$

If we like, we can replace $A_{mn}\dfrac{m!\ n!}{(m+n+1)!}$ by B_{mn}, which somewhat simplifies the last three formulae.

While these formulae are still easily manageable, the estimation of our unknowns p and h becomes increasingly cumbersome, since with increasing number of terms in the power series development for A the degree of the equations (4.10) and (4.11) in h and p increases. Solution for p and h must be done by trial and error just as much as in the case of modern methods; it will depend on the actual shape of the frequency distribution observed which method is quicker.

A second approximation will only contain terms B_{10} and B_{01} and appears relatively simple :

$$Y = \frac{1}{ph^2}\left(B_{00} + B_{10}\frac{1}{ph} + B_{01}\frac{1}{h}\right) \qquad \ldots \quad (4.12)$$

$$K = \frac{1}{ph^3}\left(\frac{1}{3}B_{00} + \frac{1}{4}B_{10}\frac{1}{ph} + \frac{1}{2}B_{01}\frac{1}{h}\right) \qquad \ldots \quad (4.13)$$

$$V = \frac{1}{p^2h^3}\left(\frac{1}{3}B_{00} + \frac{1}{2}B_{10}\frac{1}{ph} + \frac{1}{4}B_{01}\frac{1}{h}\right). \qquad \ldots \quad (4.14)$$

We can now find first approximations to the changes in p and h due to the appearance of small terms with B_{10} and B_{01} by differentiating with regard to the latter two parameters :

$$\left.\begin{array}{ll} \dfrac{\partial p}{\partial B_{10}} = \dfrac{1}{2}\dfrac{1}{hB_{00}} & \dfrac{\partial p}{\partial B_{01}} = -\dfrac{1}{2}\dfrac{p}{hB_{00}} \\[2ex] \dfrac{\partial h}{\partial B_{10}} = 0 & \dfrac{\partial h}{\partial B_{01}} = \dfrac{1}{2}\dfrac{1}{B_{00}} \end{array}\right\} \qquad \ldots \quad (4.15)$$

We are thus able to estimate the deviations of h and p from the values found in (2.4) and (2.5) for small deviations from the constant density hypothesis; B_{00} in formula (4.15) is equal to $\frac{1}{2}A$ used in Section 2.

5. RANDOM DEVIATIONS IN DISTRIBUTION

Observed frequency distributions of project coefficients will show "irregular" deviations from most of the frequency distributions so far used. It is plausible that such deviations will often be close to random deviations. From the first members of equations (2.2) and (2.3) we see that any deviations $D(\varphi, \varkappa)$ that the frequency distribution might show from one of the distributions studies will not change the quantities of capital or foreign exchange needed for the execution of a programme based on a given value of p and h if these deviations satisfy the conditions :

$$\iint_{\Delta} \varkappa D(\varphi, \varkappa) d\varphi \, d\varkappa = 0 \qquad \ldots \quad (5.1)$$

$$\iint_{\Delta} \varphi D(\varphi, \varkappa) d\varphi \, d\varkappa = 0. \qquad \ldots \quad (5.2)$$

These can be said to express that the deviations are not correlated, within the area covered by the integrals, with either $\varkappa$, the capital-output ratio or φ, the foreign currency output ratio.

THE TWENTIETH CENTURY TREND IN EMPLOYMENT IN MANUFACTURE IN INDIA—AS ILLUSTRATED BY THE CASE OF WEST BENGAL

By ALICE and DANIEL THORNER

Ecole Pratique des Hautes Etudes, Sorbonne, Paris

SUMMARY

A great expansion of manufacture in West Bengal during the three decades from 1931 to 1961 has involved, particularly in the most recent years, the installation of advanced modern facilities in a number of key industries. In absolute terms West Bengal's male working force in manufacture has grown impressively but the ratio of man employed in manufacture to the total male labour force appears to have levelled off at 15 in 100. As of 1961, this percentage was four points higher than the average for India; it was exceeded in no major State.

For the rest of this century we may expect massive decennial increments to India's total labour force. During the same years technological developments will presumably continue to raise productivity per worker and accordingly to limit the labour requirements of new manufacturing units. Substantial industrial development in terms of capital outlay and additions to productive capacity may conceivably continue for several decades with little or no increase in the percentage of the working force engaged in manufacture.

Economists who have studied the cases of England, the United States and other pioneer industrial countries have naturally tended to assume that the Indian employment statistics would eventually exhibit the classic Western shift in the percentage of workers out of agriculture and into manufacture. But the setting is entirely different: India's industrialization is taking place in the age of automation. We shall have to work out new standards for evaluating the significance of late-twentieth-century trends in working force distribution among agriculture, manufacture and other sectors.

1. THE CASE OF WEST BENGAL

The thirty years from 1931 to 1961 witnessed an unprecedented industrial development in West Bengal. Through an almost unbroken series of social and political crises—economic depression, civil disobedience, war, famine, communal strife, partition, the ensuing influx of refugees—mills, factories and workshops have been on the increase. In terms of employment, the number of men engaged in manufacture has nearly tripled. We take manufacture to mean all kinds of making and processing from hand-pounding of rice to fabricating of steel, including the production of gas and electricity, but excluding mining, quarrying and construction. Whereas the Census of 1931 recorded little more than half a million male workers under manufacture, the Census of 1961 counted more than one million and a half.

During these three decades the total population of the districts now constituting West Bengal almost doubled. The male working force in all occupations rose from five to ten millions. Since employment in manufacture expanded half again as fast as total employment, the percentage of workers in manufacture rose from 10.4 in 1931 to 15.4 in 1961.

As we can see in Table 1, which also gives comparable figures for other branches of the economy, the absolute number of men engaged in manufacture increased in each decade from 1931 to 1961. By contrast, as we note from Table 2, which presents the same data in relative terms, the whole of the percentage increase appears to have taken place between 1931 and 1941. That is, the percentage stood at 15 from 1941 to 1961.

The sharpness of the rise in the ratio of workers in manufacture to all male workers from 1931 to 1941 may result in part from the tentative nature of the basic data for 1941. The tabulations rest on a heroic reconstruction for the areas included in West Bengal after the 1947 partition of a one-in-fifty sample of the original 1941 census slips. These, it must be added, had not everywhere been properly extracted and had then been stored for fifteen years in various places under varied conditions. The huge proportion of workers whose activities were insufficiently described to permit classification (almost one of every six) suggests caution in the utilization of the absolute as well as the relative 1941 figures for the various industry groups. Nonetheless, we have no reason to doubt the general upward trend of the actual number of men at work in manufacture from 1931 to 1941 and from 1941 to 1951.

We are on somewhat firmer ground in dealing with the decade 1951 to 1961. Here also, however, our 1951 figure for manufacture includes an element of estimation. The assumptions and procedures by which we distributed a number of inadequately tabulated "earning dependents" have been described in detail in an earlier article.[1] In any case, they affect less than 75 thousand of the workers we have counted under manufacture, and could, at the very most, have served to inflate the percentage from 14.5 to 15.5.

We are accordingly safe in saying that from 1951 to 1961 more than 400 thousand men were added to West Bengal's working force in manufacture, but that the ratio of workers in manufacture to all male workers remained virtually unchanged. These were, of course, the years of the first two national five-year plans. Expansion of modern industry had become a matter of public policy. The aim of the planners was to get India embarked in every essential field so as to set up a nucleus for continuing economic growth. In West Bengal during the planning period facilities have been added or extended in engineering, chemicals, iron and steel, automobiles, and electrical machinery. To a very considerable degree, these are large-scale, heavy, capital-intensive operations. That is to say, they employ relatively little labour in relation to the investment they require, as compared, for example, with traditional handicrafts. Under these circumstances, the growth in numbers of the industrial labour force is no longer an adequate measure of the extent of industrial development.

Over these same years, 1951 to 1961, the population of West Bengal grew from 25 to 35 millions. About one-sixth of this increase resulted from the 1956 transfer from Bihar to West Bengal of an almost exclusively agricultural border area. For the State as a whole, the total number of male workers rose in the same proportion as the population, that is, nearly 40 per cent. It is a remarkable achievement in the face of the veritable *tidal* wave of persons available for employment that manufacturing has been able to provide jobs for the same proportion of workers in 1961 as in 1951.

---

[1] Alice Thorner, "Secular Trend in the Indian Economy", *The Economic Weekly*, Bombay, Vol. XIV, Numbers 28, 29 and 30, July 1962, pp. 1156-1165.

The ratio for West Bengal in 1961—15 employed in manufacture out of every 100 men at work—is markedly higher than that for All India, 11 out of 100. As we see from Table 3, no major State shows exactly as high a percentage as West Bengal, although the figures for Maharashtra, Gujarat, Kerala and Madras are, at first glance, almost identical. The leading position of West Bengal in 1961 comes out more sharply if we consider only those men who work in industries which are not of a household nature. Here West Bengal exceeds all other major States in absolute numbers (nearly 1.25 million men) as well as in ratio (12 of every 100 male workers). This comparison is perhaps the more significant one. Preliminary figures for a number of districts indicate that in Gujarat and Maharashtra, for example, the 1961 Census category of household industry includes a considerable element of livestock-keeping, an activity normally taken for analytic purposes together with agriculture.[2] It is in any case the non-household sector which encompasses all large- and medium-scale enterprises, and in effect all modern industry. With less than one-twelfth of the national population, West Bengal in 1961 claimed almost one-sixth of India's male working force in industry carried on outside the household framework.

West Bengal's rapid recent growth in manufacture contrasts sharply with a period of both absolute and relative stagnation from 1911 to 1931. Over these two decades population rose slowly and the male working force as counted in the Census actually declined. The number of men employed in manufacture fluctuated between 500 and 600 thousand; the percentage oscillated between 10 and 11 of every 100 workers. (See Tables 1 and 2).

2. THE TREND FOR INDIA

The two-phase aspect of the relative trend for employment in manufacture in West Bengal from 1911 to 1961 may also be observed in the national percentages for India (present boundaries). We read in Table 4 that, in the Indian Union as a whole, from 1911 to 1931 the number of manufacturing workers for every 100 men at work dropped from 9 to 8; by 1961, the figure had risen to 11. In absolute terms, the strength of the male working force in manufacture in India remained at about 7 million from 1911 through 1931, and doubled between 1931 and 1961. The upward turn in West Bengal since 1931 has been more pronounced than for the Indian Union, but the direction of change is the same. When we examine the trends which can be established for other states, however, we find several quite different patterns.

The most striking dissimilarity to West Bengal is provided by the ratios for manufacture in East Bengal (Pakistan) which were as low as 4 per 100 in 1911 and appeared to have sunk to 3 per 100 in 1951. To the west and south in Bihar and Orissa, the manufacture percentages also remained very modest (4 and 5) from 1911 through 1951; the 1961 figure finally indicates a rise to 7 per cent. Uttar Pradesh and the old Central Provinces (no longer existing as a State) provide illustrations of practically stationary trends at a slightly higher level—from 8 to 9 per cent. In Rajasthan we have a different story: here the percentages fell steadily decade by decade from an initial 13 in 1911 to 9 in 1951 and again in 1961.

---

[2] See "Population on the Land in the 1961 Census of Population", address delivered by Asok Mitra, I.C.S., Registrar General, to the Annual General Conference of the Indian Society of Agricultural Statistics, New Delhi, 30th December, 1962, p.41, Table 13.

On the other hand, we note a course of development parallel to that of West Bengal in the old Bombay Presidency and the former undivided Madras State. As in West Bengal, an absolute drop in the industrial working force from 1911 to 1931 was followed by a significant increase in the thirties and forties. In the Bombay Presidency the addition of 786 thousand male workers between 1931 and 1951 raised the percentage in manufacture from 11 to 15. For Madras an increment of 431 thousand workers brought the manufacture ratio up from 9 to 12. We are not in a position to follow the comparison up to 1961, since the old Bombay Presidency is now divided among Maharashtra, Gujarat, Madhya Pradesh and Mysore; while large areas of the former Madras State are now incorporated in Andhra Pradesh and Kerala. Among these successor states, however, we have already seen in Table 3 that in 1961 Maharashtra, Gujarat, Kerala and Madras had percentages above the national average both for all male workers in manufacture and for workers in "other than household" industries.

The great expansion of manufacture in West Bengal from 1931 to 1961 is thus not unique in the Indian scene, although it is the most striking case. Industrial growth over these three decades in Bengal, Bombay, Madras (also Delhi, Punjab, and certain areas of other States) has served to push up the national ratio of men engaged in manufacture to all male workers, despite the fact that the trend in so large a State as Uttar Pradesh remains at a dead level. The finding that the percentage figure for West Bengal has remained at about 15 since 1941 is particularly significant with regard to our expectations for the future trends in India as a whole.

As we have observed, the absolute increase in West Bengal's industrial labour force has been very substantial, but the same has been true for the population and accordingly for the total number of persons who must find employment in one sector or another. Moreover, the newer types of equipment which have been installed as part of India's plan to establish a core of modern industry have characteristically a high productivity, which means that their labour requirements are low.

Yet it is precisely this kind of equipment which India must continue to install if she is to emerge from economic "backwardness" into the age of automation. Only by pushing forward to higher levels in industry, science and technology can India hope to develop a modern mentality among her people and to revolutionize the productivity of her agriculture.

Following the pioneer study carried out by Colin Clark in the 1930's economists have got into the habit of conceiving economic development statistically as the shift in the *percentage* distribution of the labour force away from agriculture toward industry and services. This pattern has been well documented for such countries as England, France, Germany, Japan, and the United States. It is only natural that scholars and government planners should have expected that, as countries like India in their turn embarked upon the course of industrialization, the working force statistics would conform to the same trend. From this viewpoint, the levelling off in the percentage of the working force engaged in manufacture while the percentage in agriculture remains high is the very antithesis of economic development. So long as we remain within this framework of reference, India's current prospect would be such as to cause dismay.

In point of fact, India's current development is taking place in a manner quite different from that which characterized the industrial growth of Western Europe and North America, and in a quite different setting. India as a late-comer to the process of industrialization can take advantage of the scientific and technological progress of Europe. Productivity per worker in recently installed manufacturing units in India is, accordingly, very much greater than was possible one hundred and fifty years ago in the early stages of the industrial revolution. There are also important differences of scale. For example, the population of India today exceeds by more than forty times that of England and Wales in 1800.

What we must realize is that the economic trends of the nineteenth century in Western Europe and North America are no longer adequate guides to the path lying before India in the 1970's, 80's and 90's. Technological developments in production seem to render obsolete our standards for evaluating the significance of changes of working force distribution. We shall have to work out new criteria valid for the age of automation.

As the development in West Bengal has demonstrated, India is well started along the road to becoming a strong industrial power, and shows many signs of continuing in this direction. But from 1961 onward, India's progress toward modernization is unlikely to be reflected in a sizeable increase in the *percentage* of workers engaged in manufacture.

TABLE I. WEST BENGAL WORKING FORCE DISTRIBUTION BY INDUSTRY, 1911-1961

	1911	1921	1931	1941	1951	1961
		(males)				
total population	8,829,118	8,780,938	9,558,126	11,493,300	13,345,441	18,599,144
working force	5,549,315	5,352,978	4,999,095	4,846,600	7,220,091	10,040,230
agriculture	3,659,603	3,431,345	3,071,823	2,094,200	3,735,756	—
fishing	55,119	38,651	36,877	34,400	42,224	5,777,282
mining	40,280	38,245	24,959	57,500	85,804	—
construction	57,926	45,756	31,300	49,200	101,882	146,625
manufacture	542,778	609,369	520,962	740,600	1,119,277	1,543,521
transport	256,981	183,797	154,400	211,500	353,449	386,873
trade	324,559	323,697	332,611	405,900	778,704	836,365
public force and administration	65,497	63,352	62,854	98,100	158,173	1,349,573
professions and liberal Arts	100,332	95,255	102,254	136,600	202,903	
domestic service	173,409	182,211	200,679	149,900	206,148	
other services	61,233	62,327	58,026	85,100	121,601	
insufficiently described	211,598	278,973	402,350	783,600	314,170	

Sources : Census Tables for 1911, 1921, 1931 and 1951, *Y* sample data for 1941 from Census Paper No. 4 of 1956 ; Census of 1961, Paper No. 1 of 1962.

Note : See Appendix for discussion of sources and methods.

TABLE 2. WEST BENGAL: PERCENTAGE DISTRIBUTION OF MALE WORKING FORCE BY INDUSTRY GROUPS 1911-1961

	1911	1921	1931	1941	1951	1961
all workers	100.0	100.0	100.0	100.0	100.0	100.0
agriculture, fishing and mining	67.7	65.5	62.7	45.1	53.5	57.5
construction	1.0	1.0	0.6	1.0	1.4	1.5
manufacture	9.8	11.4	10.4	15.3	15.5	15.4
transport	4.6	3.4	3.1	4.4	4.9	3.9
trade	5.8	6.0	6.7	8.4	10.8	8.3
services	7.2	7.5	8.5	9.7	9.5	13.4
general labour and insufficiently described	3.8	5.2	8.0	16.2	4.3	

Sources : See Table 1.

TABLE 3. MALE WORKERS ENGAGED IN MANUFACTURE: SELECTED STATES, 1961

	male workers engaged in manufacture					
	absolute numbers in 000			per 100 male workers		
states	total	household industry	other industry	total	household industry	other industry
INDIA	14,534	7,366	7,168	11.3	5.7	5.6
West Bengal	1,544	300	1,244	15.4	3.0	12.4
Maharashtra	1,752	560	1,192	15.0	4.8	10.2
Gujarat	843	343	500	14.8	6.0	8.8
Kerala	582	189	393	14.7	4.8	10.0
Madras	1,432	081	751	14.2	6.7	7.4

Source : Census of India, Paper No. 1 of 1962—pp. 395, 397, 408, 409, 412, 413.
Note : See Appendix for definition of household industry.

TABLE 4. WORKERS IN MANUFACTURE PER 100 MALE WORKERS SELECTED STATES, 1911-1961

	1911	1921	1931	1941	1951	1961
INDIAN UNION	9.2	8.7	8.4	—	9.9	11.3
West Bengal	9.8	11.4	10.4	15.3	15.5	15.4
East Bengal (Pakistan)	4.2	4.6	4.1	—	3.4	—
Bihar and Orissa	5.2	4.8	4.6	4.5	4.4	7.4
Uttar Pradesh	8.7	7.7	8.0	7.4	8.3	8.9
Central Provinces	8.6	8.2	7.6	7.5	8.9	—
Rajasthan	12.6	11.1	10.2	—	8.6	8.9
Bombay Presidency	11.8	11.8	11.2	—	14.8	—
Undivided Madras	9.5	8.2	8.9	9.3	11.8	—

Sources : Census of India, 1911-1961; Census of Pakistan, 1951.
Note . See Appendix for geographical details on each State.

Appendix

The figures in Tables 1, 2, 3 and 4 have all been calculated from decennial census returns. We have used the data for males only since the reports for women in India, as elsewhere, are much less reliable. For West Bengal in particular the number of women counted as workers has always been small. From one and one third million in 1911, the figure has increased only to one and one half million by 1961. Since the female population doubled over the half century the proportion of workers reported among all women has dropped from one of every six to less than one of every ten.

In "working force" we have included all men recorded as having an occupation except for those whose activities could clearly be identified as non-productive, for example : "house proprietors", "non-cultivating landlords", and "beggars". For 1931 we have excluded from the working force all persons listed as "working dependents" in "Domestic Service". It appears that, in certain states, a misleading census instruction led to the counting under this rubric of persons (mostly women) engaged in housekeeping in their own homes. For West Bengal the number of men subtracted for this reason was 15,597, all in Burdwan District.

In the 1951 data for East Bengal which is taken from the Census of Pakistan, the total working force upon which the percentage for manufacture has been calculated excludes armed forces and non-citizens, but includes the unemployed.

The "total population" of West Bengal in the years 1911 through 1931 has been obtained by subtracting from the total for Bengal the divisions and districts cut off in 1947 to form East Bengal. In the case of the four districts which were divided between India and Pakistan we approximated the exact position by treating the whole of Malda and Jalpaiguri districts as part of West Bengal, assigning the whole of Dinajpur to East Bengal, and halving the figures for Nadia. The number of men in each of the industrial categories was calculated in the same way. In preparing entries for 1941 we utilized the Y sample tabulations both for total population and for occupational details. These tabulations were, as we have explained, unavoidably incomplete;' they do not cover Bankura, Midnapur, West Dinajpur and Cooch Bihar, slips for which were found missing or defective. Sikkim is included in the West Bengal figures for 1911, 1921 and 1931, but not subsequently. The total number of men censused in Sikkim in those years, however, never reached 60,000. For 1951 and 1961, the population of West Bengal has been taken directly from the census volumes. Thus, the 1951 figures have *not* been adjusted to include the inhabitants of the areas transferred to West Bengal from Bihar in 1956, and now constituting the Purulia District.

The industrial categories are the standard ones commonly used in historical studies of labour force trends. Agriculture includes plantations, animal husbandry, forestry and hunting. Mining includes quarrying. Trade includes banking and insurance as well as all wholesale and retail commerce. But craftsmen listed in the census volumes as "makers and sellers" of clay pots or sandals have been counted under manufacture. "Other services" are hotelkeeping, laundering, barbering and scavenging.

We have used the rubric "insufficiently described" for workers who could not be assigned to a more precise category on the basis of the description given. The greatest number of these in each year were listed as "general labourers" ; these workers were almost certainly engaged at least for part of the time as agricultural labourers. The problems involved in the treatment of this group have been discussed in the article by one of us already cited. This article also explains the procedure by which the numbers of general labourers and insufficiently described were estimated for 1951. For the purposes of the present study, we may simply say it is very unlikely that any important number of the "insufficiently described" in any year were in fact engaged in manufacture.

"Household industry" was defined in the Census of 1961 to comprise industries operated on a household scale in which all or most of the work was done by the family members. In urban areas only those manufacture or processing activities were to be included which were carried on in the home itself; the household industry in rural areas might be located anywhere within the village.

The 1911, 1921 and 1931 figures for the Indian Union in Table 4 were calculated by subtracting from the All-India totals for those years the data for the districts which were subsequently separated off to form Burma and Pakistan. The East Bengal figures for 1911, 1921 and 1931 include Sylhet District, then part of Assam. Bihar and Orissa figures for 1911, 1921 and 1931 do not include Ganjam District which was then part of Madras. The 1941 Bihar and Orissa figures exclude the Mayurbhanj and Patna States. No territorial changes of any importance affect the series for Uttar Pradesh (formerly United Provinces); the 1941 data exclude Benares State (not Benares District). Rajasthan (formerly Rajputana) includes Ajmer only in 1961. Central Provinces includes Berar in all years. This area constituted Madhya Pradesh in 1951, but should not be confused with the present State of that name. Bombay Presidency refers to the old British Districts of Bombay plus the former feudatories; these covered the whole of the Kathiawad peninsula and were scattered through Gujarat, Konkan, the Bombay Deccan and the Bombay Karnatak. Baroda has been included; Sind, Khairpur State and Aden excluded. In 1951 this area comprised the States of Bombay, Saurashtra and Kutch. The percentages for Undivided Madras include Ganjam District up through 1931 but not subsequently. The absolute increase of 430,000 in the male working force in manufacture in Madras from 1931 to 1951 has been calculated without Ganjam.

The analysis of census data on which this article rests was begun in Bombay in the years 1958-1960 under the auspices of the Indian Statistical Institute, and has been subsequently continued in Paris in the Sixth Section of the Ecole Pratique des Hautes Etudes of the Sorbonne.

A LETTER REPORT TO PROFESSOR P. C. MAHALANOBIS

By H. O. A. WOLD

University Institute of Statistics, Uppsala, Sweden

Dear Prasanta :

Fourteen years have passed since you invited me to leave the Nordic winter for a sojourn at your institute and a wide tour through India. This privilege with its immediate benefits and excitements, and its enduring and precious after-effects, induce me to give my contribution to the Festschrift Mahalanobis a personal touch. The rhapsodic form of a letter is so much the more appropriate as Radhakrishna has kindly suggested that I should report on my studies into the rationale of non-experimental analysis, with emphasis on the approach of *eo ipso* predictors.

My letter report will briefly review two groups of problems,

(i) the choice of regression, and

(ii) the stochastic rationale of interdependent (ID) versus causal chain (CC) systems,

problems that have been subject to decades of controversial and partly confused debate, group (i) especially in the 1920's and 1930's, (ii) in the 1940's and 1950's. The upshot of my review is that the problems under (i)-(ii) are closely related, and that no deep-lying difficulties are involved. What is required is rather a combination of simple arguments that have to be put together in a specific way to make sense, like the pieces of a puzzle. A key piece in the puzzle is the notion of *eo ipso* predictor. It so happens that I came across this key piece at a late stage of my studies in the area, thereby confirming and sharpening my earlier conclusions, and obtaining a unified view of the situation. This does not mean that all basic problems of the area have been solved, but it does mean that those obscure questions about (i)-(ii) that have been in the back of my mind for such long time no longer exist.

I am aware that the gradual pushing forward of the argument in a sequence of some twenty published contributions has not made for an easy grasp of my results, and several changes in terminology add to the trouble, so it is no wonder if my opposition has sometimes been misunderstood. Hence I very much welcome this opportunity to clarify my line of approach.

The limited purpose of this letter report is reflected in the references, which focus on my own contributions. A systematic review of related work would of course require a much longer exposition.

1. *Eo ipso* PREDICTORS[1]

The relation

$$y = f(x_1, \cdots, x_h) \qquad \ldots \quad (1)$$

where $f(\cdot)$ is an ordinary function, is called *deterministic*. The relation

$$y = f(x_1, \ldots, x_h) + v \qquad \ldots \quad (2)$$

where $f(\cdot)$ is an ordinary function and v a random variate, is called *stochastic*. If (2) satisfies

$$E(y \mid x_1, \ldots, x_h) = f(x_1, \ldots, x_h) \qquad \ldots \quad (3)$$

the function $f(\cdot)$ is called an *eo ipso* predictor of y. We shall mainly be concerned with relations and predictors that are linear, say

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_h x_h + v \qquad \ldots \quad (4)$$

with

$$E(y \mid x_1, \ldots, x_h) = \beta_0 + \beta_1 x_1 + \ldots + \beta_h x_h. \qquad \ldots \quad (5)$$

As to the residual variate v we note two modes of interpretation. One is to regard (2) as a deterministic relation (1) that is blurred by a residual disturbance v which sums up the effect of variates $x_{h+1}, x_{h+2}, \ldots$ that have not been explicitly taken into account. The other, not necessarily different from the first, is to disregard the possibility of deterministic relations, and to interpret $f(x_1, \ldots, x_h)$ as a representation for y that holds true only on the average, in the sense of (3).

1.1. *A first comment.* The notion of *eo ipso* predictor is of relevance in situations where the intended use of (2)-(3) is to predict y when $x_1, \ldots, x_h$ are known. In such situations the transition from deterministic relations (1) to *eo ipso* predictors (3) is a scientific evolution with important implications. Briefly stated, procedures of operation with deterministic relations are not always applicable to *eo ipso* predictors. In the first place we shall consider the operations of *inversion*, *substitution*, and *least squares estimation* of linear predictors.

1.2. *Inversion.* A deterministic relation

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_h x_h \qquad \ldots \quad (6)$$

can be reversed with respect to x_i as soon as $\beta_i \neq 0$, say

$$x_1 = \frac{1}{\beta_1} (y - \beta_0 - \beta_2 x_2 - \ldots - \beta_h x_h). \qquad \ldots \quad (7)$$

The *eo ipso* predictor (5), on the other hand, cannot be reversed. In fact, unless v is zero with probability 1 we have, as is a well known feature in correlation theory,

$$E(x_1 \mid y, x_2, \ldots, x_h) \neq \frac{1}{\beta_1} (y - \beta_0 - \beta_2 x_2 - \ldots - \beta_h x_h) \qquad \ldots \quad (8)$$

---

[1] For proofs and further comments, see Wold (1961b; 1963a; 1963b), also, (1959b; 1960; 1961a). The terminology has been changed from *unbiased* to *eo ipso* predictors on the consideration that the underlying notion of conditional expectation usually enters as an assumption, not as an implication.

1.3. *Substitution.* Given two deterministic relations (6) and

$$x_a = \gamma_a+\gamma_{a+1}x_{a+1}+\gamma_{a+2}x_{a+2}+\ldots+\gamma_h x_h \qquad \ldots \quad (9)$$

substitution of x_a into (6) gives

$$\begin{aligned} y = {} & \beta_1 x_1+\beta_2 x_2+\ldots+\beta_{a-1}x_{a-1} \\ & +\beta_0+\beta_a\gamma_a+(\beta_{a+1}+\beta_a\gamma_{a+1})x_{a+1}+\ldots+(\beta_h+\beta_a\gamma_h)x_h. \end{aligned} \qquad \ldots \quad (10)$$

When $a = 1$ the relation (10) extends to *eo ipso* predictors; that is, the two relations (5) and

$$E(x_1 \,|\, x_2, \ldots, x_h) = \gamma_1+\gamma_2 x_2+\gamma_3 x_3+\ldots+\gamma_h x_h. \qquad \ldots \quad (11)$$

imply

$$E(y \,|\, x_2, \ldots, x_h) = \beta_0+\beta_1\gamma_1+(\beta_2+\beta_1\gamma_2)x_2+\ldots+(\beta_h+\beta_1\gamma_h)x_h. \qquad \ldots \quad (12)$$

If $a > 1$, however, the corresponding parallel to (10) will in general not be valid.

1.4. *Least squares estimation.* Let

$$y = b_0+b_1x_1+\ldots+b_hx_h+u \qquad \ldots \quad (13)$$

denote relation (4) as estimated by the least squares regression of y on $x_1, \ldots, x_h$. Then the parameter estimates will under very general conditions be asymptotically consistent; in symbols,

$$\text{prob lim } b_i = \beta_1 ; \qquad i = 0, 1, \ldots, h. \qquad \ldots \quad (14)$$

The requisite conditions are (a) that the joint theoretical distribution of $x_1, \ldots, x_h$ is linearly nonsingular, and (b) that the observed moments of first and second order have the corresponding theoretical moments for probability limits as the sample size increases indefinitely.

1.5. *Equivalence with regression analysis.* As is clear from 1.4 the notion of *eo ipso* predictor is closely related to the classical notion of regression relation.

The specific feature of *eo ipso* predictors is that nothing more than (3) is assumed, or (5) in the linear case. Hence the approach of *eo ipso* predictors provides a uniform treatment of experimental data (stimulus-response experiments where each of the stimulus variables $x_1, \ldots, x_h$ is given as a nonstochastic sequence of numerical values) and nonexperimental data (the observations of $y, x_1, \ldots, x_h$ are regarded as a sample from a joint probability distribution). The approach of *eo ipso* predictors is distribution-free with regard to the residuals v. Regression analysis usually specifies the distribution of the residuals, and thereby makes it possible to obtain more detailed results, such as exact distributions for the parameter estimates b_i in (13) and (14).

In regression analysis there is a parting of the ways between the treatment of experimental and nonexperimental situations.[3] The specification

$$E(v) = 0; \quad E(v \cdot x_1) = \ldots = E(v \cdot x_h) = 0 \qquad \ldots \quad (15)$$

often enters as part of the basic assumptions of nonexperimental regression analysis. In experimental regression analysis the totality of the basic assumptions usually imply (5) as an immediate corollary. On the other hand it is easily verified that the specification (5) of linear *eo ipso* predictors always implies (15) as a simple corollary.

---

[3] H. Cramér (1945, 46, Chapters 23 and 37) gives a theoretical treatment, with illustrations and comments that emphasize the distinction at issue.

2. CHOICE OF REGRESSION : DEMAND ANALYSIS[3]

Vivid in my memory as of yesterday is my Indian winter 1949-50 and all it brought of new insights, new vistas, during my travels from Darjeeling in the north to Cape Comorin in the south, and during the months of studies and lecturing at your institute. The turmoil of new experiences was something of a shock that changed my outlook on many things, and it is tempting to expand on some of these issues. but I must stick to my subject.

You asked me to lecture on (a) time series analysis, (b) the analysis economic demand. A key theme under (b) was the choice of regression. As late as 1938, the monumental work of Henry Schultz had left this controversial problem open, inasmuch as Schultz reported two estimates of each price elasticity, one based on the regression of demand on price (and other variables), the other on the regression of price on demand, usually with a substantial difference between the two numerical values. The traditional arguments behind this dualism are the mathematical equivalence between the demand function and its inverse, say

$$d = D(p) \qquad p = D^{-1}(d) \qquad \dots \quad (16)$$

and the symmetry between the two regressions at issue.[4] I had never been able to understand this symmetry as anything else than a formal aspect. My approach was to disregard the symmetry on the subject matter argument that consumer demand is causally dependent upon market price, consumer income, and other influencing factors, and that the demand function should therefore be estimated as the regression of demand upon price, income, and other explanatory variables.[5] This causal interpretation of the demand function has further implications, two of which will be noted.

(i) *An aggregation test for the choices of regression.* A simple theoretical deduction shows that the aggregate demand for a group of commodities has an income elasticity that is the weighted average of the income elasticities of the various subgroups; testing whether this aggregation formula is valid for empirical estimates of income elasticities, it turns out that it is well satisfied by the estimates based on the regression of demand on income (price, and other explanatory variables), but not by the estimates based on other regressions, for example the orthogonal regression.

(ii) *The combined use of time series and cross section data.* According to the causal interpretation of demand functions, the income elasticity should be the same whether it is estimated from time series data or from cross sections, for in principle the two estimates measure the response to income changes of one and the same group of consumers. For one thing, this argument can be exploited for comparative checks on the income elasticities derived from the two types of data. For another thing, the causal argument establishes the device of *conditional regression analysis*: When price and income elasticities are estimated from time series data the regression is often marred by collinearity between income, price and other explanatory variables; the collinearity is reduced or avoided by taking the income elasticity to be known, as given by an estimate based on cross section data.

---

[3] For an instructive review of how the problems pose themselves and their treatment develops up to the end of the 1930's, see G. Stigler (1954).

[4] The symmetry argument has been strongly emphasized by R. Frisch (1928, 1934). An early reference is E. P. Mackeprang (1906).

[5] See, also for the subsequent points (i)-(ii), Wold (1940; 1942; 1943-44; 1945).

May I say how much I was benefited from the opportunity the Indian lectures and seminars gave me to sum up and discuss my research endeavours in demand analysis. In addition, my plans to write a monograph on the subject began to take concrete form, in Calcutta and during the ten weeks' tour with guest lectures and seminars at seven or eight universities, and I brought home a rather detailed layout of the volume (Wold and Juréen, 1952). As to the choice of regression, the monograph emphasizes that the choice between different estimation techniques must be decided upon from case to case with regard to the theoretical model that underlies the application at issue. Once the model is specified the choice of estimation procedure is a technical matter that in point of principle is noncontroversial.

The measurement of demand functions is treated in the monograph as an application of regression analysis procedures. The treatment in terms of *eo ipso* predictors came into the picture much later (Wold, 1959b; 1960; 1961a;1961b; 1963a; 1963b). The specification of the demand functions in terms of *eo ipso* predictors is completely in line with the above arguments. In conclusion, I shall comment upon how naturally the notion of *eo ipso* predictors links up with earlier arguments, making for clarification and sharpening the conclusions.

2.1. By its very definition, (3) or (5), it is clear that the approach of *eo ipso* predictors is appropriate in applied work that involves the prediction of y when $x_1, \ldots, x_h$ is given. Operational arguments of this type are of old standing in regression analysis, of course, and the notion of *eo ipso* predictors links up these arguments with the purely mathematical notion of conditional expectation.[6]

2.2. The causal interpretation of demand functions reveals the formalism of the alleged symmetry between the two regressions of demand upon price (or income) and of price (or income) upon demand. According to economic theory, demand is causally influenced by price and income; hence between price (or income) and demand there is the same asymmetry as between stimulus and response in a controlled experiment. The notion of *eo ipso* predictor links up this argument with the asymmetry between the variables in a conditional expectation $E(y \mid x)$, the variable x taking an arbitrarily specified value, whereas y stands for an average or expected value.

2.3. For the fundamental operative procedures with predictive relations considered in comments 1.1-4 we have seen that their theoretical rationale is quite simple, thanks to the specification in terms of *eo ipso* predictors. Moreover, as compared with the traditional specification in regression analysis, the rationale can be established under somewhat more general assumptions; notably, this is so for least squares estimation (see Wold, 1963b). The same holds true for operative procedures of more special nature, cases in point being the aggregate test and the device of conditional regression analysis referred to earlier in this section.

---

[6] Specific reference is made to L. Hurwicz (1950), whose approach is related to but less general than that of *eo ipso* predictors.

3. MULTIRELATION MODEL BUILDING : INTERDEPENDENT (ID) *Versus* CAUSAL CHAIN (CC) SYSTEMS[7]

Reference is made to R. A. Fisher's general distinction between the three main steps of a statistical study (Fisher, 1925):

(i) Specification of hypotheses;

(ii) Estimation of parameters;

(iii) Testing of hypotheses.

The debate on "the choice of regression" in the 1920's and 1930's was essentially concerned with the statistical treatment of unirelation models. In hindsight, the long debate reflects a most fundamental transition in the specification of hypotheses, namely from deterministic to stochastic models. In unirelation models the key issue of the transition is the operation of inversion, which is trivial and perfectly permissible in deterministic models, but not so in models that take the disturbances into explicit account.

Turning to the debate on "the rationale of ID versus CC systems" in the 1940's and 1950's the scene changes from unirelation to multirelation model building. Again, the debate reflects the transition from deterministic to stochastic models, now with the complication that there are two key operations in play, namely inversion and substitution, which are trivial in deterministic models but do not always extend to stochastic models. And again, as we shall see, the obscure issues dissolve in the light of the notion of *eo ipso* predictors.

3.1. When Trygve Haavelmo (1943) initiated what is now known as the approach of interdependent (ID) systems, he opened up a vigorous attack on least squares regression as an estimation technique for multirelation models, dismissing the least squares approach without qualification. The argument at issue thus started as a matter of estimation techniques. With reference to the pioneering macroeconomic models constructed by Jan Tinbergen for the Netherlands (1937) and US (1939), models of the type now known as recursive (RC) or causal chain (CC) systems,[8] R. Bentzel and I (1946) showed that least squares regression under fairly general assumptions does apply to CC systems. To the best of our knowledge, our 1946 paper is the first where CC and ID systems are compared, or recognized as distinct types of model, or even referred to in one and the same paper. The upshot of our paper was that there are deepgoing differences between the two types of model, and that the choice between them is not a matter of estimation technique.

---

[7] The readers of this report are assumed to have some orientation in the literature on ID and CC systems. For text-book treatments, with references to earlier work, see G. Tintner (1952) and L. Klein (1953).

[8] The two terms are usually taken to be synonymous. It has been proposed (Wold, 1961b) to use the term RC system when referring to the stochastic properties of the model, and CC system when the subject-matter content of the model has been specified, but in the present brief review I do not insist on this distinction

Then for some twelve years, from 1947 through 1958, I explored several aspects of the situation. Specifically : (a) It was emphasized that CC systems, thanks to the subdiagonal design of the coefficient matrix of the explanatory current endogenous variables, allow a straightforward causal interpretation; (Wold, 1949, 1954). (b) It was shown that any given set of time series can, to any prescribed accuracy, be formally represented as a CC system with the required noncorrelation assumptions of type (15)(Wold, 1948; 1951). (c) Evidence was given that the difference in approach between CC and ID systems is of considerable numerical relevance. Part of the evidence under (c) was drawn from a case study where an ID system was compared with a CC system as applied to the same set of statistical data; (Wold, 1959a).

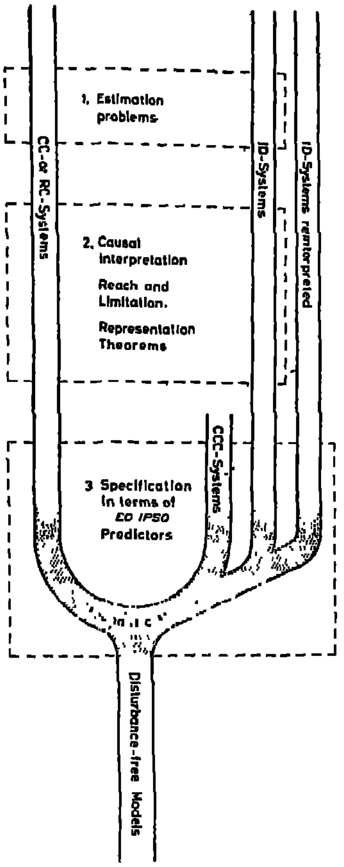


Fig. 1. The puzzle of ID and CC systems : The consecutive exploration of three main pieces.

With gradually increasing emphasis I pointed out the attractive features of CC systems on the one hand, and voiced scepticism about the interdependent approach on the other, referring in particular to the obscurity around the question as to what precise causal interpretation could be given to the coefficients of ID systems. To this phase belongs a review presented at the European meeting of Econometric Society in Uppsala in 1954, (Wold, 1955), and also the first part of the triptych paper by R. Strotz and myself (1960) in the Frisch Festschrift of *Econometrica*.

The unsatisfactory state of things in applied work with ID systems was emphasized by Haavelmo in his Presidential address to Econometric Society (1958). The basic divergence of views among protagonists of ID systems is amply demonstrated by the contributions to a recent symposium in *Econometrica* (October, 1960).

By this time I had gradually become convinced about the following views : (a) CC systems are of very general scope for purposes of dynamic model building, not only in econometrics, but in nonexperimental situations in general; (b) CC and ID systems are innovations, not only in econometrics, but in scientific model building in general; (c) an essential element of the innovation lies in the explicit allowance for residual errors, and it is this innovation that has been the cause of so much confused and partly controversial debate. As indicated in the figure the dualism between CC and ID systems disappears in the special case when the models are disturbance-free (deterministic).

My contribution (1959b) to the Harald Cramér volume was planned from the point of view briefly indicated in the previous paragraph. Early in this project it dawned upon me that it would be clarifying to explore the rationale of uni-and multi-relation model building with regard to the general procedures of operation with relations specified in terms of conditional expectations, or to use the new name, in terms of *eo ipso* predictors. As to operations of fundamental relevance in unirelation model building, the situation has been briefly reviewed in the comments 1.1-2 and 1.4. Coming to operations with multirelation models, we note, as a first comment :

Remark: The comments 1.1-2 and 1.4 are valid for any *eo ipso* predictor, whether it constitutes a unirelation model or forms part of a multirelation model. Comment 1.3 refers only to multirelation models.

3.2. *Forecasting by the chain principle.*[9] For a brief review of the rationale of CC versus ID systems from the point of view of *eo ipso* predictors, we shall consider three types of forecasting models where the forecasts are derived by the *chain principle,* that is, the forecasts for a sequence of variables or vectors y_t, y_{t+1}, ... are obtained recursively, the forecast for y_{t+k} being used as an actual observation when calculating the forecast for y_{t+k+1}.

Matrix notation will be used, writing $y_t = [y_{1t}, \dots, y_{mt}]$ for the vector of current endogenous variables, z_t for the vector of lagged endogenous variables y and current or lagged exogenous variables x, and ε_t or other Greek letters for the vector of residuals. Time t runs through integral values $t = 0, \pm 1, \pm 2, \dots$. All variables will be measured as deviations from their mean.

(A) *Vector regression* :

$$y_t = Rz_t + \varepsilon_t \text{ with } E(y_t \mid z_t) = R z_t. \qquad \dots \quad (17a\text{-}b)$$

In words, system (17a) gives the current endogenous vector y_t explicitly in terms of the predetermined vector z_t. Otherwise expressed, the primary (or structural) form of the system coincides with the reduced form.

Since forecasting by the chain principle is a matter of iterated substitutions, system (17) must for all t be specified in accordance with the condition $a = 1$ in (9)-(12) lest the forecasts be biased. At first sight this might seem to be a very restrictive condition. Nonetheless, model (17) is in principle of general scope: Given any set of endogenous variables y_t and exogenous variables x_t, the parameters R can be specified so that the requisite conditions be fulfilled to any prescribed accuracy. The general scope of model (17) is a feature closely related to the representation theorem known as *predictive decomposition of* stationary time series.[10]

(B) *Causal chain (CC) systems :*

$$y_t = Ay_t + Bz_t + v_t \text{ with } E(y_t \mid \hat{y}_t, z_t) = Ay_t + Bz_t \qquad \dots \quad (18a\text{-}b)$$

---

[9] Here I draw from Wold (1963a). See also (1959b); (1961b); and Wold-Jureen (1952), Chapter 2.7, especially as to the causal aspects of models A-C.

[10] References Wold (1938; 1948; 1951) and Wold-Jureen (1952, Chapter 12,5–7).

where the matrix $A = [a_{ik}]$ is subdiagonal, i.e. $a_{ik} = 0$ for $k \geqslant i$, and the cap over $\hat{y}_t$ indicates that as a conditional vector for a variable y_{it} the vector $\hat{y}_t$ does not involve y_{it}. The reduced form of model (18) may be written as in (17a-b); to specify,

$$y_t = R\,z_t + \varepsilon_t \text{ with } E(y_t | z_t) = Rz_t \qquad \dots \quad (19a\text{-}b)$$

where

$$R = (I-A)^{-1} . B; \quad \varepsilon_t = (I-A)^{-1} . v_t . \qquad \dots \quad (20a\text{-}b)$$

System (18) has a more general pattern of parameters than (17), and accordingly gives the model builder greater freedom in constructing the system. An innovating feature of the model is that its primary form (18) differs from the reduced form (19). The general scope of the model will be noted : Given any set of endogenous and exogenous variables, the parameters R can be specified so that to any prescribed accuracy (a) all relations both of the primary form (18) and the reduced form (19) make *eo ipso* predictors ; (b) conditions for application of the chain principle are fulfilled.[11]

(C) *Interdependent (ID) systems :*

$$\tilde{y} = C'y_t + Dz_t + \omega_t \qquad \dots \quad (21a\text{-}b)$$

where the vector $\tilde{y}_t = [y_{i_1t}, y_{i_2t}, \dots, y_{i_mt}] = \tilde{I}y_t$ is composed of some set of current endogenous variables (some y_{it} may be missing in y_t, and others may appear twice or more), and C' is arbitrary except that $c_{ki_k} = 0$. Hence the k-th relation in (21) may be written

$$y_{i_kt} = \Sigma' c_{kj} y_{jt} + \sum_j d_{kj} z_{jt} + \omega_{kt} \qquad (22)$$

where the prime in Σ' indicates that the term $c_{ki_k}\, y_{i_kt}$ is missing. Again, the reduced form may be written as in (17a-b), this time with

$$R = (\tilde{I} - C')^{-1} . D ; \quad \varepsilon_t = (I - C')^{-1} . \omega_t . \qquad \dots \quad (23a\text{-}b)$$

The relations of the reduced form are specified as *eo ipso* predictors,

$$E(y_t | z_t) = Rz_t . \qquad \dots \quad (24)$$

In system (21) the parameter pattern is still more general than in (18). The ensuing greater freedom in the model construction is however brought at a price: Between the primary form (20) and the reduced form (21) there emerges a profound dualism, inasmuch as it will in general not be possible to specify the relations of both forms as *eo ipso* predictors; that is, (24) in general implies

$$E(y_{i_kt} | \hat{y}_t, z_t) \neq \sum_j{}' c_{kj} y_{jt} + \sum_j d_{kj} z_{jt} . \qquad \dots \quad (25)$$

3.3. *Levels of generalization. A reinterpretation of ID systems.* Having seen that models A–C represent three levels of increasing generalization, the situation may be summed up as follows : In models A–C the primary forms (17), (18), (21) are specified as behaviour relations, and possibly they also involve a number of identities. In model A each relation makes an *eo ipso* predictor (and thereby, as noted in 2.2 each behaviour relation has the formal appearance of a cause-effect relation), and the model as it stands provides forecasts according to the chain principle, period by period. The same statements hold true for model B, that is, for CC systems, with the qualification that the forecasting proceeds chainwise from one variable to the next within each period and then chainwise from one period to the next. In model C, that is, in ID systems, the specification of the behaviour relations as *eo ipso* predictors breaks down, and thereby the cause-effect specification of the behaviour relations; the forecasting proceeds chainwise from period to period on the basis of the reduced form, and within the periods there is no chainwise forecasting from one variable to the next.[12]

---

[11] See the references in footnote 10. E. A. Robinson (1963) develops the subject from the point of view of minimum-delay structure, a fundamental notion belonging to information theory.

In line with the causal specification, any behaviour relation in a CC system can be used as an autonomous forecasting relation, the pre-requisite values of the explanatory variables being obtained by ancillary forecasts, but not so in ID systems. For example, let

$$d_t = \alpha+\beta p_t+v_t \text{ with } E(d_t|p_t) = \alpha+\beta p_t \qquad \ldots \quad (26\text{a-b})$$

be a demand relation, specified as an *eo ipso* predictor. The relation (26a-b) is perfectly admissible in a CC system, but in general not in an ID system, and in the CC system it can be used separately for forecasting d_t on the basis of ancillary forecasts for price p_t. Or otherwise expressed, suppose the variables in (26) are logarithmic; then if (26) is part of a CC system β is the elasticity of demand with respect to price, but if (26) belongs to an ID system this interpretation in general is not valid.

The obscurity around the cause-effect interpretation of the behaviour relations of ID systems has been the main source of the scepticism that I have voiced for such long time against the approach. Now quite recently—and much to my dismay for not having seen this before —it has turned out that the obscurity can be removed by a slight change in the specification of ID systems.[13] To set forth this point, let us modify the ID system (21) as follows, to be called the corresponding ID* system,

$$y_t = C'y_t^*+Dz_t+\omega_t^*. \qquad \ldots \quad (27)$$

Here the matrices C' and D are numerically the same as in (21), the predetermined variables z_t are also the same, whereas the current endogenous variables in the right-hand member are replaced by their expected values as given by the reduced form,

$$y_t^* = E(y_t|z_t). \qquad \ldots \quad (28)$$

As can readily be shown, we then have

$$E(y_t|y_t^*, z_t) = C'y_t^*+Dz_t. \qquad \ldots \quad (29)$$

As to the reduced form, this is always the same for an ID system and the corresponding ID* system.

Thanks to (29), the behaviour relations of ID* systems allow the same interpretation in terms of *eo ipso* predictors and cause-effect relationships as in CC systems. The difference relative to ordinary ID systems is that the current endogenous variables do not enter as explanatory variables and stimulus variables, but instead their expected values as given by (28). Again with reference to the simple demand relation (26a) for illustration this time assuming that it is part of an ID system, so that (26b) is not valid, the demand relation of the corresponding ID* system reads

---

[12] Knowing that in ID systems the reduced form but in general not the primary form can be specified in terms of *eo ipso* predictors, it is instructive to consider the opposite possibility of specifying the coefficient matrices C' and D so that the relations of the primary form, but in general not of the reduced form, will make *eo ipso* predictors. I have called this type of model *conditional causal chain* (CCC) systems (Wold, 1959b; 1960) and in the same connection I have used the term *pure causal chain* (PCC) systems for the models (18) that in current usage, as well as in this report, are called RC or CC systems.

[13] Wold (1963a, Theorem 10). Also Wold (1961a, Remark 3.2.2b).

$$d_t = \alpha+\beta p_t^*+\omega_t^* \text{ with } E(d_t|p_t^*) = \alpha+\beta . p_t^* \qquad \dots \quad (30a\text{-}b)$$

where

$$\omega_t^* = \omega_t + \beta(p_t - p_t^*) \qquad \dots \quad (31)$$

and p_t^* is given by the residual-free part of the reduced form. In the case of logarithmic variables, β is the demand elasticity with respect to expected or predicted price p_t^*.

The parameters of an ID* system are numerically the same as for the corresponding ID system. Hence the problem of parameter estimation is precisely the same for ID* as for ID systems, and meets the same formidable difficulties. Among the estimation techniques developed for ID systems, specific reference is made to H. Theil's (1958) two-stage method of least squares, inasmuch as its theoretical rationale is particularly well suited for an extension to ID* systems.

ID* systems thus constitute a logically consistent type of model, more general than CC systems. To mark the difference from the customary specification of ID systems, the ID*-systems might be called *expectional interdependent systems.*

Final comments. In referring to my studies into the rationale of multirelation model building, some writers have quoted as a statement of mine that the real world, the actual course of events, is a CC system. This is a gross misunderstanding of my position. CC systems are models, constructs, and as such they may be more or less realistic, but they must not be confused with the real world. It is quite another matter that any given set of time-series can, under very general assumptions, and to any prescribed accuracy, be formally represented as a CC system. Hence, and since CC systems can be estimated by ordinary least-squares regression, it is quite natural that model building in practice should start with vector regression of type (17) or with CC systems (18), on the argument that the possibilities of simple models should first be explored before turning to more complicated designs.

This letter focusses on the general stochastic structure of CC systems and related models. To build a CC system in practice, the formal framework of the model must be filled with subject-matter content. This task opens up an array of important considerations; problems that fall outside the scope of this report.

Jan Tinbergen's first multirelation models date from before World War II, but although this period seems distant it lies only some 25 years back. Tinbergen's approach brought a radical increase of aspiration levels in econometric model building, so from this point of view 25 years is a short time, and so much the shorter as there has been the obscure issues about ID versus CC systems to sort out, as well as other basic questions. Now even if it is early, the horizon is eagerly watched for signs of definite progress. Symptomatic in this respect is the symposium in *Econometrica*, October 1960, with the motto: "Simultaneous equation estimation — any verdict yet ?"

Well, dear Prasanta, the replies of the *Econometrica* symposium were on the whole in the negative. However, the symposium already lies a few years in the past and there are several new approaches in the picture. I should like to conclude my letter by the conjecture that the time will soon be ripe for a verdict.

Sincerely yours
Herman W.

References

Bentzel, R. and Wold, H. (1946): On statistical demand analysis from the viewpoint of simultaneous equations. *Skand. Aktuarietidskrift*, 29, 95–114.

Christ, C. *et al* (1960): Symposium on simultaneous equations estimation. *Econometrica*, 28, 835–871.

Cramér, H. (1945): *Mathematical Methods of Statistics*, Almqvist and Wiksell, Uppsala. Princeton University Press, Princeton. N. J., 1946.

Fisher, R. A. (1925): *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh, 14th ed. 1954.

Frisch, R. (1928): Correlation and scatter in statistical variables. *Nordisk Statistisk Tidskrift*, 8, 36–102.

——— (1934): Statistical confluence analysis by means of complete regression systems. *University Institute of Economics, Oslo*, Publication No. 5.

Haavelmo, T. (1943): The statistical implications of a system of simultaneous equations. *Econometrica*, 11, 1–12.

——— (1958): The role of the econometrician in the advancement of economic theory. *Econometrica*, 26, 351–357.

Hurwicz, L. (1950): Prediction and least squares. *Statistical Inference in Dynamic Economic Models*, edited by T C. Koopmans, Wiley, New York, 266–300.

Klein, L. R. (1953): *A Textbook of Econometrics*, Row Peterson, Evanston, Illinois.

Mackeprang, E. P. (1906): *Pristeorier*, Bagge, Copenhagen.

Robinson, E. A. (1963): Structural properties of stationary stochastic processes with applications. *Time Series Analysis Symposium*, edited by M. Rosenblatt, Wiley, New York, 170–192.

Robinson, E. A. and Wold, H. (1963): Appendix. Minimum-delay structure of least-square and *co-ipso* predicting systems for stationary stochastic processes. *Time Series Analysis Symposium*, edited by M. Rosenblatt, Wiley, New York, 192–196.

Schultz, H. (1938): *The Theory and Measurement of Demand*, University of Chicago Press, Chicago, Ill.

Stigler, G. (1954): The early history of empirical studies of consumer behaviour. *Journal of Political Economy*, 62, 95–113.

Strotz, H. and Wold, H. (1960): Recursive vs. nonrecursive systems: An attempt at synthesis. *Econometrica*, 28, 417–427.

Theil, H. (1958): *Economic Forecasts and Policy*, North-Holland Publ. Co., Amsterdam.

Tinbergen, J. (1937): *An Econometric Approach to Business Cycle Problems*, Hermann, Paris.

——— (1939): *Statistical Testing of Business Cycle Theories*, I: *A Method and its Application to Investment Activity*. II · *Business Cycles in the United States of America*, 1919–1932. League of Nations, Geneva.

Tintner, G. (1952): *Econometrics*, Wiley, New York.

Wold, H. (1938): *A Study in the Analysis of Stationary Time Series*, Almqvist and Wiksell, Uppsala, 2nd ed. 1954.

——— (1940): The demand for agricultural products and its sensitivity to price and income changes. (Swedish). *Statens Offentliga Utredningar*, 16.

——— (1942): Eine schwedische Untersuchung über die Elastizität der Nachfrage. *Metron*, 14, 263–285.

——— (1943-44): A synthesis of pure demand analysis, I–III. *Skand. Aktuarietidskrift*, 26, 85–118 and 220–263; 27, 69–120.

——— (1945): A theorem on regression coefficients obtained from successively extended sets of variables. *Skand. Aktuarietidskrift*, 28, 181–200.

——— (1948): On prediction in stationary time-series. *Ann. Math. Stat.*, 19, 558–567.

——— (1949): Statistical estimation of economic relationships. *Econometrica*, 17, Supplement, 1–22.

——— (1951): Dynamic systems of the recursive type—Economic and statistical aspects. *Sankhyā*, 11, 205–216.

——— (1954): Causality and econometrics. *Econometrica*, 22, 162–177.

——— (1955): Possibilités et limitations des systêmes à chaîne causale. *Cahiers Sém. Econ. R. Roy*, No. 3, 81–101.

——— (1956): Causal inference from observational data. A review of ends and means. *J. Roy. Stat. Soc*, Ser. A., 119, 28–60.

——— (1959a): A case study of interdependent versus causal chain systems. *Rev. Internat. Statist. Inst.*, 26, 5–25.

——— (1959b): Ends and means of econometric model building. Basic considerations reviewed. *Probability and Statistics. The Harald Cramér Volume*, edited by U. Grenander, Almqvist and Wiksell; Wiley, New York, 1960, 355–434.

——— (1960): A generalization of causal chain models. *Econometrica*, 28, 443–463.

——— (1961a): Construction principles of simultaneous equations models in econometrics *Bulletin de l'Institut International de Statitique*, 38, No. 4, 111–138.

——— (1961b): Unbiased predictors. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1, University of California Press, 719–761.

——— (1963a): Forecasting by the chain principle. *Time Series Analysis Symposium*, edited by M. Rosenblatt, Wiley, New York, 471–497.

——— (1963b): On the consistency of least squares prediction. *Sankhyā*, Series A, 25, 211–216.

Wold, H. and Juréen, L. (1952): *Demand Analysis. A Study in Econometrics*, Almqvist and Wiksell, Stockholm; Wiley, New York, 1953.

SCIENTIFIC CONTRIBUTIONS OF PROFESSOR P. C. MAHALANOBIS

(Figures within [ ] refer to the serial numbers of papers listed at the end)

It is not easy to make a proper assessment of the scientific work of Professor P. C. Mahalanobis as his contributions have highly individual and even unique qualities. He did not consider statistics as a narrow subject confined to the mathematical theory of probability or routine analysis of data in applied research or collection of facts as an aid to administrative decisions. He has, indeed, made signal contributions to each of these aspects of statistics. But he took a wider view of statistics as a new technology for increasing the efficiency of human efforts in the widest sense. This has naturally aroused his interest in various fields of scientific enquiry and applied problems and enabled him to enrich the science of statistics with a practical base of great depth and spread.

1. LEADER OF STATISTICAL MOVEMENT IN INDIA

To narrate the achievements of Professor Mahalanobis in the field of statistics would mean narrating the recent history of statistical development in India. Statistics as a separate discipline was recognised by the scientific world even in the most advanced countries of the world only at the turn of the present century, and in India it was practically unknown in the first quarter of the century. There was no teaching of or research in statistics at the universities, no statistical society or journal of statistics. The official statistics were collected only as a by-product of administration. A sudden spurt of statistical activity occurred about 30 years ago which was to put India, in the course of a few years, not far from the centre of the world statistical map. The Indian Statistical Institute with facilities for high-level research, training and execution of large scale project work was founded by Mahalanobis in 1931. '*Sankhyā*', the Indian Journal of Statistics was started in 1933, of which 25 volumes have been published and of which he remains editor to this day. Sample surveys on a large scale covering whole provinces were started on behalf of Government under the leadership of the Institute from 1937. The first Indian Statistical Conference was held in 1938 under the Chairmanship of the late Sir Ronald Fisher. Post-graduate courses in statistics were opened for the first time in India in the Calcutta University in 1941 under the guidance of Mahalanobis; similar courses were opened in other universities subsequently. The Indian Science Congress decided to start a separate section for mathematics and statistics in 1942 and a separate section for Statistics in 1945. A Central Statistical Unit was established by the Government of India in 1949, to work under the technical guidance of Professor Mahalanobis as Honorary Statistical Adviser to the Cabinet; two years later the Central Statistical Organisation was started and still later a Department of Statistics was created which is still working under Mahalanobis's guidance. The National Sample Survey was established in 1950 for the collection of socio-economic data through sample surveys on a continuing basis, a unique feature in a national statistical system. The Indian Statistical Institute was recognised as an institution of national importance and empowered to award degrees in statistics by a Central Act in 1959.

This phenomenal progress was due to the foresight, leadership and efforts of Professor Mahalanobis. As the founder-director of the Indian Statistical Institute he has been taking the initiative in expanding the Institute's scope of activities and in increasing the field of application of statistical methods in many directions. The movement which started in 1932 was itself the result of the deep conviction which the Professor had developed, during the preceding ten years of advanced studies and research in various fields, about the importance of statistics in an under-developed country like India.

Professor Mahalanobis was drawn into statistics by a series of random events. After taking his honours degree in India, Mahalanobis joined King's College, Cambridge University, in 1913 for higher studies and research; he took Part I of Mathematical Tripos in 1914 and Part II of the Physics Tripos, with a first class in 1915. He returned to India in 1915 with the intention of going back to Cambridge for research in physics for which he had been awarded a senior scholarship by King's College.

Just before his departure from Cambridge, his tutor, W. H. Macaulay, had quite casually drawn his attention to the journal *Biometrika* and *Biometric Tables* edited by Karl Pearson, copies of which Mahalanobis brought with him to India; studying these volumes he got seriously interested in statistics. Soon after his return to India he accepted the post of a professor of physics at the Presidency College, Calcutta, and decided to stay in India and carry on his statistical studies alongside his work of teaching physics. A little later in 1917, Sir Brajendra Nath Seal (then University Professor of Philosophy) had asked Mahalanobis to help in the work of a committee, appointed by the Calcutta University, to enquire into the examination system ; this offered good opportunities of using statistical methods for the analysis of examination data. Mahalanobis also got other problems, usually of a practical nature (relative efficiency of different types of vaccines, effect of weather conditions on incidence of diseases, inter-caste marriages etc.) on which several notes and reports were prepared.

Speaking about the early activities of Mahalanobis the late Sir Ronald Fisher had observed—

> "In those early days Professor Mahalanobis was a Professor of Physics, but was no ordinary Professor of Physics. Instead of doing his duty and teaching his students, only that science and technology which has led to the atomic bomb, he thought that he would serve the fellow-men by opening the door to knowledge, obtained by statistical means, to knowledge of the natural world and knowledge of the world of society in which we all find ourselves. Perhaps, he thought, as many of us think, that most of the follies and crimes of the Government are due to ignorance—ignorance of the people they serve, and the ignorance of the people whom they are all too willing to fight."

2. MAHALANOBIS DISTANCE

For some time Mahalanobis carried on his statistical studies in his own home with the help of part-time computers engaged by him. A chance meeting with N. Annandale (then Director of the Zoological and Anthropological Survey of India) at the Nagpur Session of the Indian Science Congress in 1920 led to Mahalanobis's first systematic work on a statistical study of the anthropometric measurements taken by Annandale on Anglo-Indians (of mixed British and Indian parentage) in Calcutta. Mahalanobis's first paper [1] was on 'the Statistical Analysis of Anglo-Indian Stature' which was published in the *Records of the Indian Museum* in 1922. This was a remarkable paper in many ways and exhibited the abilities of Mahalanobis as a statistician, to cross examine his data, to extract the information by appropriate statistical methods and to suggest further lines of investigation. Then followed a series of papers [6, 10, 11, 12, 13, 14, 23, 37, 58, 59, 62, 66, 76, 78, 81, 100, 129] on the theory and application of multivariate analysis to anthropological data the most important of which is [6] the 'Analysis of race mixture in Bengal', the Presidential address delivered to the Anthropological Section of the Science Congress.

The anthropometric studies led to the formulation of D^2-statistics, known in statistical literature as *Mahalanobis distance*, which has proved to be a valuable tool not only in taxonomy but in many other fields including economics and geology. He had also discussed the problem of statistical taxonomy from a logical point of view in 1937, and had laid down the fundamental axioms (hypotheses) under which a stable classification of taxa could be obtained [62]. A rich field of research in multivariate analysis opened up, to which some of his colleagues at the Institute made valuable contributions.

3. METEOROLOGICAL RESEARCH

In accordance with the usual practice, Mahalanobis's paper on statistical analysis of the stature of Angle-Indians had been referred to Sir Gilbert Walker (then Director General of Observatories) for opinion which led to Sir Gilbert asking Mahalanobis to undertake a statistical study of some meteorological problems. This resulted in two memoirs [2, 3] and a note [4] on upper air variates being published in 1923. Statistical deviations from the average were for the first time explicitly separated from experimental errors of observations. This led, by purely statistical methods, to the location, at the height of about 4 kilometers above sea level, of the region of highest control for changes in meteorological conditions on the surface of the earth (a result which was rediscovered several years later by Franz Bauer in Germany from physical considerations).

The meteorological studies led to Mahalanobis being appointed Meteorologist in Calcutta, in charge of the eastern region; he worked in this post for more than 3 years from 1922 to 1926 in addition to his duties as professor of physics at the Presidency College. During this period he started some work on the statistical analysis of meteorological data; in one study he found that there was greater rainfall in Calcutta in July during the darker phase of the moon and that the difference was statistically significant ; these results were published several years later [67].

4. Reconstruction of information through scrutiny

Most of the statistical methods available in the early twenties took into account only the variability of measurements (i.e. of individuals) in a population but not of the additional variation due to errors of measurement. These errors may be of a random nature inherent in any method of obtaining a scale reading or of a systematic type introducing a bias, or due to wrong recording. Such errors, if not recognised and allowed for in statistical analysis, may lead to wrong inferences. Mahalanobis seems to have been the first to point out and study the effects of measurement and recording errors on statistical parameters. In his memoir [3] on the analysis of meteorological data of the upper air published in 1923 he worked out the effects of errors of measurement on the estimates of the correlation coefficient and standard deviation.

He developed a highly scientific method of scrutinising data for recording and systematic errors, and reconstructed a large series of anthropometric measurements published by Risley, which had been earlier condemned as faulty and unsuitable for statistical analysis being full of errors [23, 37]. This is a difficult problem in extraction or reconstruction of information, which is indeed more of an art in which Mahalanobis is a great master. He has an uncanny sense for figures and can often point out at a glance mistakes in any series of figures coming under his scrutiny. The publication entitled, 'Revision of Risley's anthropometric data' in the first number of *Sankhyā* in 1931 brought a letter of appreciation from the late Sir Ronald Fisher.

5. Early examples of operations research

In 1922 a disastrous flood had occurred in North Bengal. An expert committee of engineers appointed by Government were about to recommend the construction of expensive retarding basins to hold up the flood water when the question was referred to Mahalanobis for examination. A statistical study of rainfall and floods extending over a period of about 50 years showed that the proposed retarding basins would be of no value in controlling floods in North Bengal. Floods were occurring in the area because of obstructions to the outflow of the excess rainfall by river and railways systems without adequate bridges ; the real need was improvement of rapid drainage and not holding up the flood water. Specific remedies were recommended, many of which were implemented and proved effective [9].

A similar question of flood control in Orissa was referred to Mahalanobis, after a severe flood in the Brahmini river in 1926. This resulted in a detailed study of the areal and time distribution of rainfall in relation to floods in Orissa rivers. An expert committee of engineers were of opinion that a catastrophic rise had occurred in the river-bed of the Brahmani, and they recommended increasing the height of river embankments by several feet. The statistical study covering a period of about 60 years showed that no change had occurred in the river bed, and showed that the construction of dams for holding up of excessive flood water in the upper reaches of the rivers was likely to be of value in this region [15, 16, 83]. Mahalanobis pointed out in 1928 that the real need was to construct dams in the upper reaches of the rivers to hold up the water for purposes of flood control and generation of electric power. He also gave the first calculations for a multi-purpose (flood control, irrigation and power) scheme for the Mahanadi system in Orissa, which formed the basis of the Hirakud hydroelectric project which was inaugurated about thirty years later in 1957. Two big reports were submitted incorporating the extensive studies of rainfall and floods in Orissa [15, 16]. Some of the results in the reports on rainfall and floods were published in a summary form in a paper [83].

In retrospect one finds that Mahalanobis's early work on river floods is in the nature of Operations Research which has been introduced as a separate discipline after the second world war. Another example of such an approach is the study he made in a note [61] submitted to the Government of Bengal in 1937 on the Hooghly-Howrah flushing and irrigation scheme. He had estimated the expected increase in yield of paddy by the introduction of an irrigation scheme by which high floods in the Damodar river were intended to be used to supplement the rainfall. This report was of help in developing in later years the Damodar Valley multi-purpose hydroelectric project.

6. Educational tests

Mahalanobis developed an interest also in the construction and performance of educational tests in his early days of statistical research. Group tests in the medium of the Bengali language were constructed and used in the Institute [22, 36, 39]. Studies were also made of the correlation between intelligence or aptitude tests and success in school leaving certificate and other examinations, and of standardisation of tests [42, 57, 63].

7. Errors in field experimentation

In 1925 Mahalanobis had the occasion to study the probable errors of results of agricultural experiments in which six varieties of paddy had been sown in a systematic manner in ten replicated blocks. In the resulting paper [7] the variation in soil heterogeneity was sought to be eliminated by fitting non-linear graduating curves (a method which was used independently by J. Neyman several years later). In complete ignorance of the work on the design of experiments done a little earlier by R. A. Fisher, Mahalanobis suggested in this paper that each strip or plot should be divided into a number of sub-plots with a view to securing replicated observations for the computation of error. This paper placed him in touch with R. A. Fisher, and Mahalanobis immediately began to study seriously Fisherian methods of field experiments. During a visit to England in 1926, he met Fisher and close personal relations were established. The friendship between Mahalanobis and Fisher was strengthened by Fisher's frequent visits to the Indian Statistical Institute, specially during the last ten years. They had common views on the foundations and methodological aspects of statistics and also on statistics as a new technology of the present age. Sir Ronald's death in 1962 removed a very close friend of Mahalanobis.

Mahalanobis did valuable service in introducing in India as early as 1929 Fisherian techniques in conducting field experiments and analysis of data. A number of papers by himself and in collaboration with his colleague S. S. Bose have appeared during 1932 and 1940 on field experimentation and related statistical methods [17-21, 24, 34, 38, 40, 43, 44, 46-51, 53, 54, 65, 68, 71, 80, 82, 89].

8. SAMPLE SURVEYS, SEQUENTIAL SAMPLING AND INTERPENETRATING NETWORK OF SAMPLES (IPNS)

Then came the epoch making investigations on the technique of large scale sample surveys, with which Professor Mahalanobis's name will always be associatod. Systematic work on survey of agricultural crops began in 1937, which culminated in a large scale survey of the acreage and yield of Jute crop in 1941, covering the whole province of Bengal, and was extended to all important crops in both Bengal and Bihar in 1943. This was followed by sample surveys for collecting socio-economic data, public preference etc. [69-75, 79, 84-88, 90-99]. These demonstrated the possibility of using large-scale sample surveys to collect information quickly and economically with sufficient accuracy for most practical purposes. Three notable contributions were made to the technique, namely, the concept of optimum designs of surveys (to give the highest precision at a given cost or the minimum cost for an assigned level of precision), pilot surveys and interpenetrating sub-samples. The pilot surveys represent a very general approach of which a significant example is sequential sampling developed 10 years later by the late Abraham Wald, and provide a systematic method for progressively improving the design of the survey, utilising the previous information on cost and variance functions. The interpenetrating network of samples, used in a proper way, is the only possible method to estimate non-sampling errors which are inevitable in large-scale sample surveys and which are sometimes of a larger order of magnitude than sampling errors; also the estimation of the non-sampling errors is necessary for assessing the precision of a final estimate. The first full-length treatment of these concepts is contained in the memoir on large-scale sample surveys which appeared in 1944 in the *Philosophical Transactions* of the Royal Society [116].

Speaking about the originality of ideas which Mahalanobis was trying out in the development of sample designs for estimation of Jute crop, the late Sir Ronald mentioned :—

> "What was striking therefore in the Indian contribution to the problems of sample survey, was that it combined a clear realization of statistical principles with a down-to-earth experience of the practical difficulties, in a country in which education was, on the whole, backward, of getting such work done, on a large scale, and yet with the primary scientific requirement of ascertainable and demonstrable precision. From that point the standard set has never gone back, and publicly organised censuses and surveys in all parts of the world look to the Indian Sample Survey as a basis for comparison. Difficulties, of course, there are, in India, and also elsewhere, but the speed and economy, as well as the accuracy of sample surveys when competently organised, have now been put beyond question."

A large number of surveys were conducted from 1937 to 1950 in Bengal for collecting information on crops and socio-economic data which gave opportunities for improving the design of sample surveys and for gaining experience in the collection of data from the field [102-115, 118-120, 122-127, 132-135, 138, 139, 141]. Time was ripe for launching nationwide sample surveys. Several hurdles had to be crossed in convincing the administrators that gaps in national statistics could be filled through data obtained by sampling and there would be a need for continuous collection of information to assess the progress of economic development and to make policy decisions. Mahalanobis won the battle and in 1950 the National Sample Survey was established. It is a continuing survey, in which information is collected year by year with the help of a whole-time field organisation, spread all over India, and provides periodic estimates on social and economic factors affecting the national economy [142, 146, 176, 179].

9. NATIONAL PLANNING

In 1954, at the request of Prime Minister Nehru and the Planning Commission of India, Mahalanobis undertook a study to find ways and means of increasing national income at a reasonably rapid rate and at the same time solving the problem of unemployment if possible in 10 years. The problem of national development had been engaging the attention of Mahalanobis even earlier. In his Presidential Address, 'Why Statistics'? to the Science Congress in 1950 [131] and in two lectures [140, 143] in 1952 he gave some calculations emphasizing the need of a high rate of investments in heavy industries. Elaborating on his earlier ideas [144, 148] Mahalanobis developed econometric models (known as Mahalanobis's two and four sector models) for determining optimum investments in different sectors of the national economy [149-154, 156-165] and he submitted to Government the Draft Plan-frame for the Second Five Year Plan of India in 1955. Mahalanobis's findings supplied a rational strategy for national development and formed the basis of the Second and Third Five Year Plans of India. He is continuing his interest in the planned economic development of India as a Member of the Planning Commission of India. In recent years he has taken special interest in requirements of technical and scientific manpower and the need of establishing the scientific base for a rapid economic and social progress of India [166-170, 172, 174, 175, 177, 178, 180, 181].

10. FRACTILE GRAPHICAL ANALYSIS

It is said that often a scientist's best work is done in his young age and as he grows old he can only elaborate on his earlier ideas. But Mahalanobis seems to be an exception in this respect. During his long years of research he has been continually providing new ideas and new tools which would form permanent additions to statistical techniques. To the generalised distance and axioms of taxonomy, pilot surveys, inter-penetrating sub-samples (IPNS), concepts of randomness of space fields, and economic growth models, he has recently (1958) added the new and powerful tool of Fractile Graphical Analysis (FGA) (158).

This novel approach was developed out of a necessity to compare the socio-economic conditions of groups of people, differing in place and time, using the data collected in the various rounds of the National Sample Survey. The method is simple, extremely versatile with a very wide field of applications, and can be modified to meet various needs. The method is graphical and is based on a geometrical concept of error which enables a study of the relationship between two variates, and can also provide a generalised measure of the separation or the generalised difference between two different "universes" of study. Considerable research has been done on theoretical aspects and a good deal of experience has been gained in the use of F.G.A., in various fields (like economics, demography, psychometry, botany, biometry, physical measurements etc.) during the last few years and there is now a sizeable literature on the subject [173, 180].

EDITORS

BIBLIOGRAPHY

LIST OF SCIENTIFIC PAPERS CONTRIBUTED BY PROFESSOR P. C. MAHALANOBIS

1922

1. Statistical analysis of Anglo-Indian stature. *Rec. Ind. Museum*, **23**, 1-96.

1923

2. On the seat of activity in the upper air. *Memoirs Ind. Met. Dept.*, **24**, 1-9.
3. On errors of observation and upper air relationship. *Memoirs Ind. Met. Dept.*, **24**, 11-19.
4. Correlation of upper air variables. *Nature*, **112**, 323-324.
5. Statistical note on the significant character of local variation in the proportion of dextral and sinistral shells in samples of the snail Buliminus Dextro Sinister from the salt range, Punjab. *Rec. Ind. Museum*, **25**, 399-403.

1925

6. Analysis of race-mixture in Bengal. Presidential Address, Anthropological Section, Indian Science Congress. *Jour. and Proc. Asiat. Soc. Bengal*, **23**, 301-333.
7. Probable error of experiments in agriculture. *Agr. Jour. Ind.*, **20**, 96.

1926

8. Note on seasonal variations in the incidence of diarrhoea, dysentery and appendicitis in Calcutta. *Cal. Med. Jour.*, October, 1926.

1927

9. Report on rainfall and flood in North Bengal during the period 1870-1922. (In 2 volumes with 29 tables and 28 maps), submitted to the Government of Bengal, 1-90.

1928

10. Statistical study of the Chinese head. *Man in India*, **8**, 107-122.
11. On the need for standardization in measurements on the living. *Biometrika*, **20**, 1-31.

1930

12. On tests and measures of group divergence, Part 1 : Theoretical formulae. *Jour. and Proc. Asiat. Soc. Bengal,* New Series, **26,** 541-588.

13. Statistical study of certain anthropometric measurements from Sweden. *Biometrika,* **22,** 94-108.

1931

14. Statistical analysis of Anglo-Indian head length. *Rec. Ind. Museum,* **23,** 97-149.

15. Statistical study of the level of the rivers of Orissa and the rainfall in the catchment areas during the period 1868-1928. Submitted to the Government of Bihar and Orissa.

16. (and Chakravarti, N.) : Statistical report on river floods in Orissa during the period 1868-1928. (In 20 chapters with 47 tables and 27 maps), submitted to the Government of Bihar and Orissa.

1932

17. Statistical note on the method of comparing mean values based on small samples. *Ind. Jour. Agr. Sc.,* **2,** 28-41.

18. Statistical note on certain rice breeding experiments in the Central Provinces. *Ind. Jour. Agr. Sc.,* **2,** 157-169.

19. Auxiliary tables for Fisher's Z-test in analysis of variance. (Statistical notes for agricultural workers, No. 3). *Ind. Jour. Agri. Sc.,* **2,** 679-693.

20. Rice and potato experiments at Sriniketan. (Agricultural Department of the Visva-Bharati) 1931. (Statistical notes for agricultural workers, No. 4). *Ind. Jour. Agri. Sc.,* **2,** 694-701.

21. (and Bose, S. S.) : Note on the variation of the percentage infection of Wilt disease in the cotton plant in Surat. (Statistical notes for agricultural workers, No. 5). *Ind. Jour. Agri. Sc.,* **2,** 704-709.

1933

22. Studies in educational test, No. 1. The reliability of a group test of intelligence in Bengali. With five figures. *Sankhyā,* **1,** 25-49.

23. Revision of Risley's anthropometric data relating to the tribes and castes of Bengal. *Sankhyā,* **1,** 76-105.

24. Tables for L-tests. *Sankhyā,* **1,** 109-122.

25. Effects of fertilizers on the variability of the yield and rate of shedding of buds and flowers and bolls in the cotton plant in Surat. (Statistical notes for Agricultural workers, No. 6). *Ind. Jour. Agri. Sc.,* **3,** 131-138.

26. Note on field technique. (Government of Bombay, Department of Agriculture).

27. Use of the method of paired differences for estimating the significance of field trials. (Statistical notes for agricultural workers No. 11). *Ind. Jour. Agri. Sc.*, **3**, 349-352.

28. On the need of randomization of plots in field trials. (Statistical notes for agricultural workers, No. 13). *Ind. Jour. Agri. Sc.*, **3**, 549-551.

29. Use of random sampling numbers in agricultural experiments. (Statistical notes for agricultural workers, No. 14). *Ind. Jour. Agri. Sc.*, 3, 1108-1115.

30. (and Bose, S. S.): Effect of the time of application of fertilizers on the yield and the rate of shedding of buds, flowers and bolls in the cotton plant in Surat. (Statistical notes for agricultural workers, No. 7). *Ind. Jour. Agri. Sc.*, **3**, 139-146.

31. (and Bose, S. S.): Effect of different doses of Nitrogen on the rate of shedding of buds, flowers and bolls in the cotton plant in Surat. (Statistical notes for agricultural workers, No. 8). *Ind. Jour. Agri. Sc.*, **3**, 147-154.

32. (and Bose, S. S.): Certain varietal studies on the cotton plant in Surat. (Statistical notes for agricultural workers, No. 9). *Ind. Jour. Agri. Sci.*, 3, 339-344.

33. (and Bose, S. S.): Analysis of a manurial experiment on wheat conducted at Sakrand, Sind. (Statistical notes for agricultural workers, No. 10). *Ind. Jour. Agri. Sc.*, **3**, 345-348.

34. (and Bose, S. S.): Analysis of varietal test with wheat conducted at Sakrand, Sind, 1931-32. (Statistical notes for agricultural workers, No. 12). *Ind. Jour. Agri. Sc.*, **3**, 544-548.

35. Editorial. *Sankhyā*, **1**, 1-4.

1934

36. Studies in educational tests. No. 2. The age variation of scores in a group test of intelligence in Bengali. (With eight figures.) *Sankhyā*, **1**, 231-244.

37. Revision of Risley's anthropometric data relating to the Chittagong hill tribes. *Sankhyā*, **1**, 267-276.

38. Analysis of rotational experiments with cotton, groundnut, and juar in Berar, with notes on designs for rotational experiments. (Statistical notes for agricultural workers, No. 15). *Ind. Jour. Agri. Sc.*, **4**, 361-385.

39. Studies in educational tests, No. 4. A preliminary study of the intelligence quotient of Bengali School children. (With ten figures.) *Sankhyā*, **1**, 407-426.

40. (and Bose, S. S.): Statistical note on the effect of pests on the yield of sugarcane and the quality of cane-juice. *Sankhyā*, **1**, 399-406.

41. (with the cooperation of Bose, S. S., Ray, P. R. and Banerjee, S. K.): Tables for random samples from a normal population. *Sankhyā*, **1**, 289-328.

42. (and Chakravarti, K. N.): Studies in educational tests, No. 3. Analysis of marks in the school-leaving certificate examination in the United Provinces, India, 1919. (With ten figures.) *Sankhyā*, **1**, 245-266.

1935

43. (and Bose, S. S.): On estimating individual yields in the case of mixed-up yields of two or more plots in agricultural experiments. *Science and Culture*, **1**, 205.

44. (with Bose, S. S. and Harrison, C. J.): Effect of manurial dressings, weather conditions and manufacturing processes on the quality of tea at Tocklai experimental station, Assam. *Sankhyā*, **2**, 33-42.

45. Study of consumer preference in Calcutta. (Submitted to D. J. Keymer & Co.).

1936

46. (with Bose, S. S. and Banerjee, K. C.): Studies in tiller variation. (Statistical notes for agricultural workers, No. 20). *Ind. Jour. Agri. Sc.*, **6**, 1122-1133.

47. (with Bose, S. S., and Sengupta, S. C.): Statistical analysis of a manurial experiment on Napier grass (Pennisetum Perpureum) by the method of co-variance. (Statistical notes for agricultural workers, No. 18.) *Agri. Live-stock in India*, **6**, 183-194.

48. (with Bose, S. S. and Ganguly, P. M.): Frequency distribution of plot yields and optimum size of plots in a uniformity trial with rice in Assam. (Statistical notes for agricultural workers No. 19). *Ind. Jour. Agri. Sc.*, **6**, 1107-1121.

49. (with Bose, S. S. and Chakravarti, S. C.): Effect of different methods of harvest on the estimated error of field experiments on rice. (Statistical notes for agricultural workers, No. 21.) *Agri. Live-stock in India*. **6**, 814-825.

50. (with Bose, S. S. and Chakravarti, S. C.) : Complex experiment on rice at the Chinsurah farm, Bengal, 1933-34. (Statistical notes for agricultural workers, No. 16). *Ind. Jour. Agri. Sc.*, **6**, 34-51.

51. (with Bose, S. S. and Kulkarni, R. K.): On the influence of shape and size of plots on the effective precision of field experiments with juar (Andropogon Sorghum). (Statistical notes for agricultural workers, No. 17). *Ind. Jour. Agri. Sc.*, **6**, 460-474.

52. New theory of ancient Indian chronology. *Sankhyā*, **2**, 309-320.

53. Note on cotton prices in relation to quality and yield. *Sankhyā*, **2**, 135-142.

54. Editorial note on the margin of error in the calculation of the cost of cultivation and profit. *Sankhyā*, **2**, 121-124.

55. Karl Pearson, 1857-1936. *Sankhyā*, **2**, 363-378.

56. Note on the statistical and biometric writings of Karl Pearson (with selected references). *Sankhyā*, **2**, 411-422.

57. (assisted by Roy, S.): School marks and intelligence test scores. *Sankhyā*, **2**, 397-402.

58. On the generalised distance in statistics. *Proc. Nat. Inst. Sc., India*, **2**, 49-55.

1937

59. On the accuracy of profile measurements with a photographic profiloscope. *Sankhyā*, **3**, 65-72.

60. Need of a sample survey of the growth of population in India. *Sankhyā*, **3**, 58.

61. Statistical note on the Hooghly-Howrah flushing and irrigation scheme. Submitted to the Government of Bengal, Irrigation Department, 1937, 1-44.

62. (with Bose, R. C. and Roy, S. N.) : Normalization of statistical variates and the use of rectangular co-ordinates in the theory of sampling distributions. *Sankhyā*, **3**, 1-40.

63. (and Acharya, D. P.): Note on the correlation between results in the college and university examinations. *Sankhyā*, **3**, 239-244.

64. (and Bose, S. S.): On an exact test of association between the occurrence of thunderstorm and an abnormal ionisation. *Sankhyā*, **3**, 249-252.

65. (and Bose, S. S.): On estimating individual yields in the case of mixed-up yields of two or more plots in field experiment. *Sankhyā*, **4**, 103-120.

66. (and Datta, B. N.): Note on the foot and stature correlation of certain Bengal castes and tribes. *Sankhyā*, **3**, 245-248.

67. Note on the influence of lunar phase on the rainfall in the month of July in Calcutta, 1878-1924. *Sankhyā*, **3**, 233-238.

68. Review of the application of the statistical theory to agricultural field experiments in India. Proceedings of the Second meeting of the crops and soils wing of the Board of Agriculture and Animal Husbandry, Lahore, 6 December, 1937, Govt. of India, 1-15 [1].

1938

69. Note on grid sampling. *Science and Culture*, **4**, 300.

70. Statistical report on the experimental crop census, 1937. Indian Central Jute Committee.

1939

71. (with Bose, S. S. and Khanna, K. L.): Note on the optimum shape and size of plots for sugarcane experiments in Bihar. (Statistical notes for agricultural workers, No. 24.) *Ind. Jour. Agri. Sc.*, **9**, 807-816.

72. First report on the crop census of 1938. Submitted to the Indian Central Jute Committee, 1939, 1-110.

73. Enquiry into the prevalence of drinking tea among middle class Indian families in Calcutta, 1939. Submitted to the Tea Market Expansion Board, 1939, *Sankhyā*, **6**, 283-312.

74. Progress report of the jute census scheme for 1939. Submitted to the Indian Central Jute Committee, 1939.

1940

75. Report on the sample census of jute in 1939. Submitted to the Indian Central Jute Committee, 1940, 1-146.

76. Application of statistical methods in physical anthropometry, (appears as second part of "Discussion on the application of statistical methods in anthropometry"), in *Proceedings of the Second Indian Statistical Conference*, Lahore, and *Sankhyā*, 4, 591-598.

77. Errors of observation in physical measurements. *Science and Culture*, **5**, 443–445.

78. Anthropological observation on the Anglo-Indians of Calcutta, Part III, Statistical analysis of measurements of seven characters. *Rec. Ind. Museum.*, **23**, 141,

79. Sample survey of the acreage under jute in Bengal. *Sankhyā*, **4**, 511–530.

80. Discussion on planning of experiments. *Sankhyā*, **4**, 530-531.

81. Application of statistical methods in physical anthropometry. *Sankhyā*, **4**, 594-598.

82. Review of the application of statistical theory of agricultural field experiments in India (Revised). *Ind. Jour. Agri. Sc.*, **10**, 192-236.

83. Rain storms and river floods in Orissa (with a map). *Sankhyā*, **5**, 1-20.

84. Statistical note on crop-cutting experiments on paddy in Mymensingh. Submitted to the Government of Bengal, 1940.

85. Statistical report on crop-cutting experiments on jute 1940. Submitted to the Indian Central Jute Committee, 1940.

86. Preliminary report on the sample census of the area under jute. Submitted to the Indian Central Jute Committee, 1940.

87. Note on the expenditure on tea among working class families in Howrah and Kankinara. Submitted to the Indian Tea Market Expansion Board.

88. Report on the sample census of the area under jute in Bengal in 1940. Submitted to the Indian Central Jute Committee, 1941.

89. (and Nair, K. R.): Simplified method of analysis of quasi-factorial experiments in square lattice with a preliminary note on joint analysis of yield of paddy and straw. (Statistical notes for agricultural workers, No. 25). *Ind. Jour. Agri. Sci.*, **10**, 633-685.

1941

90. Statistical survey of public opinion. *Modern Review*, **69**, 1941.

91. Note on random fields. *Science and Culture*, **7**, 54.

92. Statistical note on nutritional investigations in college hostels in Calcutta. *Sankhyā*, **5**, 439-448.

93. Statistical report on a sample survey of the prevalence of drinking tea in Nagpur City in 1940. Submitted to the Tea Market Expansion Board, March 1941.

94. Preliminary statistical report on the regional survey of Borer pests of sugarcane, December 1940–March 1941. Submitted to the Imperial Council of Agricultural Research, June 1941.

95. Statistical report on the rupee census (Revised). *Report on Currency and Finance*, 1940-41, Reserve Bank of India.

96. Statistical report on crop-cutting experiments on jute in Bengal, 1940. Submitted to the Indian Central Jute Committee, June 1941.

97. Report on programme preference and broadcast reactions, Calcutta, April–May 1941. Submitted to the Department of Information, Government of India, July, 1941.

98. Report on the sampling technique for forecasting the bark-yield of Cinchona plants: Experiment series B, 1940-41. Submitted to the Superintendent Cinchona Culture in Bengal, Mungpoo, October 1941.

99. General report on the sample census of the area under jute in Bengal, 1941. Submitted to the Indian Central Jute Committee, 1941.

100. (and Bose, C.): Correlation between anthropometric characters in some Bengal castes and tribes. *Sankhyā*, **5**, 249-260.

101. (and Nair, K. R.): Statistical analysis of experiments on differential limen values for lifted weights. *Sankhyā*, **5**, 285-294.

1942

102. Statistical definition of standard yield of crops. *Sankhyā*, **6**, 97-98.

103. Sample surveys. Presidential Address, Section of Mathematics and Statistics, Indian Science Congress, Baroda, 1942, *Proc. Twentyninth Sci. Cong;* Part II, 25-46.

104. Note on the life of a rupee note, July 1940–June 1941. Reserve Bank of India, 1942.

105. Family budget enquiries of labourers, Jagaddal. Scheme for Labour Family Budget Enquiry and Preparation of cost of Living Index of Industrial Workers. Submitted to the Board of Economic Enquiry, Bengal, January 1942.

106. Report on the precision of family budget enquiry at Jagaddal, August 1942. Submitted to the Board of Economic Enquiry, Bengal, 1942.

107. Preliminary report on Burdwan-Hooghly-Howrah crop-cutting survey. Submitted to the Government of Bengal, Irrigation Department, 1942.

108. Preliminary report on the sample census of the area under jute in Bengal, 1942. Submitted to the Government of Bengal, 1942.

109. Preliminary report on the sample census of the area under aus paddy in Bengal, 1942. Submitted to the Government of Bengal, 1942.

110. Preliminary report on the crop-cutting experiment on jute, 1942. Submitted to the Government of Bengal, 1942.

111. Enquiry into the prevalence of drinking tea among middle-class Indian families in Calcutta, 1939. *Sankhyā*, **6**, 283-312.

1943

112. Report on the regional survey of Borer pests on sugarcane, December 1940-March 1941. Submitted to the Imperial Council of Agricultural Research, 1943, 1-284.

113. Final report on the sample census of the area under jute and aus paddy in Bengal, 1943. Sample Surveys for Estimating the Acreage and yield of jute in Bengal, 1937-43. Submitted to the Government of Bengal, 1943, 1-51, (37 Tables).

114. Preliminary report on the sample census of the area under aman paddy in Bengal, 1943. Submitted to the Government of Bengal, 1943.

115. (and Ghosh, B. N.) : Statistical analysis of data relating to incidence of pests and diseases on different varieties of sugarcane. Indian Statistical Conference, Calcutta, 1943.

1944

116. On large-scale sample surveys. *Phil. Trans. Roy. Soc.*, London, Series B, **231**, 329–451.

117. Organisation of statistics in the post-war period, *Proc. Nat. Inst. Sc. India*, **10**, 69-78,

1945

118. Report on the Bihar crop survey, Rabi season, 1943-44. *Sankhyā*, **7**, 29-106.

1946

119. Recent experiments in statistical sampling in the Indian Statistical Institute (with discussion). *J. Roy. Stat. Soc.*, Series A, **109**, 325–378. Published, in book form under the title *Experiments in Statistical Sampling in the Indian Statistical Institute,* jointly by Asia Publishing House and Statistical Publishing Society.

120. Sample surveys of crop yields in India. *Sankhyā*, **7**, 269-280.

121. Distribution of the Muslims in population of India. *Sankhyā*, **7**, 429-434.

122. Traffic census on new Howrah bridge. Submitted to the Government of India, 1946.

123. (with Mukherji, R. K. and Ghose, A.): Sample survey of after-effects of the Bengal famine of 1943. *Sankhyā*, **7**, 337-400.

124. Use of small-size plots in sample surveys for crop yields. *Nature*, **158**, 798-799,

125. Bengal famine. *Asiatic Review*, October 1946.

1947

126. Enquiry into the economics of agricultural labour and rural indebtedness. Submitted to the Government of India, 1947.

127. Enquiry into the economics and statistics of road development. Submitted to the Government of India, 1947.

1948

128. Walter A. Shewhart and statistical quality control in India. *Sankhyā*, **9**, 51-60.

1949

129. (with Majumdar, D. N. and Rao, C. R.): Anthropometric survey of the United Provinces, 1941: a statistical study. *Sankhyā*, **9**, 90-324.

130. Statistical tools in resource appraisal and utilization. *Proc. U.N.S.C.U.U.R.*, **1**, 1-14.

1950

131. Why Statistics ? Presidential Address, Indian Science Congress, Thirtyseventh Session, Poona. *Sankhyā*, **10**, 195-228.

132. Age tables based on the *Y*-samples. Submitted to the Government of India, 1950.

133. Survey of rural indebtedness. Final Report: Rural Indebtedness Enquiry, 1946-47. Submitted to the Government of West Bengal, 1950.

134. Survey of economic condition of agricultural labour. Submitted to the Government of West Bengal, 1950.

135. Cost and accuracy of results in sampling and complete enumeration. *Bull. Int. Stat. Inst.*, **32**(2), 210–213.

1951

136. Role of mathematical statistics in secondary education. *Bull. Int. Stat. Inst.*, **33**(5), 323–334.

137. Professional training in statistics. *Bull. Int. Stat. Inst.*, **33**(5), 335–342.

138. Means of livelihood and industries tables based on the *Y*-sample. Submitted to the Government of India, 1951.

139. (and Sengupta, J. M.): On the size of sample cuts in crop-cutting experiments in the Indian Statistical Institute, 1939–1950. *Bull. Int. Stat. Inst.*, **33**(2), 359–404.

1952

140. Statistical methods in national development. The Thirteenth Jagadish Chandra Bose Memorial Lecture, Bose Institute, Calcutta, 30 November, 1951. *Science and Culture*, **17**, 497–504.

141. Some aspects of the design of sample surveys. *Sankhyā*, **12**, 1–7, 1952–53.

142. National Sample Survey: General Report No. 1 on the first round, October 1950—March 1951. Submitted to the Government of India, 1952.

143. National income, investment, and national development. Lecture delivered at the National Institute of Sciences of India, New Delhi, 4 October 1952.

1953

144. Some observations on the process of growth of national income. *Sankhyā*, **12**, 307–312.

1954

145. Foundations of statistics. *Dialectica,* **8**, 95–111.

146. On some aspects of the Indian national sample survey. *Bull. Int. Stat. Inst.*, **34** (2), 5–14. (Spanish translation: Sobre algunos aspectos de la encuesta nacional de la India por muestras. *Estadistica,* **55**, 260.)

147. Report on survey of saver's preference in Delhi State. Submitted to the Ministry of Finance, Government of India, 1954.

148. Studies relating to planning for national development. Address delivered on the occasion of inauguration by Prime Minister Jawaharlal Nehru of the Studies relating to Planning for National Development, at the Indian Statistical Institute, Calcutta, 3 November, 1954. Published in 1961 in book form under the title *Talks on Planning,* jointly by Asia Publishing House and Statistical Publishing Society.

1955

149. Draft plan-frame ; Recommendations for the formulation of the Second Five Year Plan 1956-57—1960-61. Submitted to the Government of India, 17 March, 1955. *Sankhyā,* **16**, 63-90. Published in 1961 in book form under the title *Talks on Planning*, jointly by Asia Publishing House and Statistical Publishing Society.

150. Approach to planning in India. Based on a talk broadcast from All India Radio, 11 September, 1955. Published in 1961 in book form under the title *Talks on Planning,* 47-54, jointly by Asia Publishing House and Statistical Publishing Society.

151. Approach of operational research to planning in India. *Sankhyā,* **16**, 3-130. Published in 1963 in book form under the title *An Approach of Operational Research to Planning in India,* jointly by Asia Publishing House and Statistical Publishing Society.

1956

152. Statistics must have purpose. Presidential address, Pakistan Statistical Conference, Lahore, February 1956.

153. Agricultural statistics in relation to planning. Address delivered at the Ninth Annual Meeting of the Indian Society of Agricultural Statistics, 7 January, 1956, *Jour. Ind. Soc. Agri. Stat.*, **8**, 5-13.

154. Address delivered at the Thirty Second Annual General meeting of the Geological, Mining and Metallurgical Society of India, Calcutta, 28 September 1956, *Quarterly Journal of the Geological, Mining and Metallurgical Society of India*, **28**, 87-88.

1957

155. University teaching of social science: Statistics, a survey. Prepared and edited on behalf of the International Statistical Institute, Paris, UNESCO.

1958

156. Enfoque de la planeacion en la India. *Trimestre Economico*, **25**, 654-663.

157. Industrializatsiya—klyuch k ukrepleniyu nezavisimosti. (Industrialization a key to the consolidation of independence), *Sovremennyi Vostok* (Contemporary East), No. **12**, 15-18.

158. Method of fractile graphical analysis with some surmises of results. *Transactions of the Bose Research Institute.*, **22**, 223-230.

159. Science and national planning. Anniversary address delivered at the National Institute of Sciences of India, New Delhi, 5 January, 1958, *Sankhyā*, **20**, 69-106, and published in 1961 in book form under the title *Talks on Planning*, jointly by Asia Publishing House and Statistical Publishing Society.

160. Some observations on the 1960 world census of agriculture. *Bull. Int. Stat. Ins.* 36(4), 214-221, (Spanish translation : Algunas observaciones sobre el censo agropecnario mundial de 1960, *Estadistica,* March, 11–31, 1960).

161. Industrialization of underdeveloped countries—a means to peace. *Bulletin of the Atomic Scientists*, 15, (12), 1959 *Sankhyā*, **22**, 173-182. Published in 1961 in book form under the title *Talks on Planning*, jointly by Asia Publishing House and Statistical Publishing Society.

1959

162. Problems of economic development in India and other underdeveloped countries in relation to world affairs. *Bulletin of the International House of Japan*, 3, 10-15.

163. Izuchenio problem industrializatii slaborazvitikh stran. (Study of the problems of industrialization in the underdeveloped countries). *Sovremennyi Vostok* (Contemporary East), No. 9, (English translation published in 1961 in book form under the title *Talks on Planning*, jointly by Asia Publishing House and Statistical Publishing Society, 137-142.

164. Review of recent developments in the organisation of science in India. Presented at the 6th General Assembly and Scientific symposium held under the auspices of the World Federation of Scientific Workers, Warsaw, 1959. *Vijnan Karmee*, **11**, 13-27.

165. Heralding a new epoch, *A Study of Nehru ;* edited by Rafiq Zakaria, 1959, 309-319, and published in book form under the title *Talks on Planning*, 1-8, jointly by Asia Publishing House and Statistical Publishing Society.

166. Next steps in planning. Anniversary address, delivered at the National Institute of Sciences of India, New Delhi, 20 January 1959, *Sankhyā*, **22**, 143-172, and published in 1961 in book form under the title *Talks on Planning*, 93-126, jointly by Asia Publishing House and Statistical Publishing Society.

167. Need of scientific and technical man-power for economic development. Based on a talk broadcast from All India Radio, 23 September 1959. Published in 1961 in book form under the title *Talks on Planning*, 143-146. jointly by Asia Publishing House and Statistical Publishing Society.

168. Unemployment. Address delivered as the sectional Chairman of the Second All India Labour Economics Conference, Agra, January 1959, *Indian Journal of Labour Economics*, **2**, 39-45, 1959, and published in 1961 in book form under the title *Talks on Planning*, 147-152, jointly by Asia Publishing House and Statistical Publishing Society.

1960

169. Incentives and scientific and technical personnel. *Dainik Samachar*, 1960.

170. Perspective planning. Address delivered at the Third Session of the SEANZA Central Banking Course in Bombay.

171. (and Mukherjee, M.) : Operational research models used for planning in India. Presented at the Second Operational Research International Congress, September, 1960.

172. Scientific workers in the UK, USA, and USSR. *Science and Culture*, **2**, 101-110.

173. Method of fractile graphical analysis. *Econometrica*, **28**, 325-351, and *Sankhyā*, Series A, **23**, 41-64.

174. Labour problems in a mixed economy. Address delivered at the Third All India Labour Economics Conference, Madras, 2 January 1960. *Indian Journal of Labour Economics*, **3**, 1-6, 1960, and published in 1961 in book form under the title *Talks on Planning*, jointly by the Asia Publishing House and Statistical Publishing Society, 153-159.

175. Note on problems of scientific personnel. Draft recommendations placed before the Scientific Personnel Committee, 24 March, 1960. *Science and Culture*, **27**, 101-128.

1961

176. (and Lahiri, D.B.): Analysis of errors in censuses and surveys with special reference to experience in India. *Bull. Int. Stat. Inst*, **38**, (2), 401-433, and *Sankhyā*, Series A, **20**, 325-358.

177. Role of science in economic and national development. Lecture delivered at the University of Sofia, 4 December, 1961. *Indian Journal of Public Administration*, **8**, 153-160.

178. Statistics for economic development. *Journal of the Operations Research Society of Japan*, **3**, 97-112,

179. *Talks on Planning*, (collected addresses and broadcasts on various aspects of planning in India). Asia Publishing House and Statistical Publishing Society.

1962

180. Preliminary note on the consumption of cereals in India. *Bull. Int. Stat. Inst.*, **39** (4), 53-76.

181. Scientific base of economic development. Presented at the Conference for International Cooperation and Partnership, Salzburg-Vienna, 1-7 July, 1962.

182. Problems of internal transformation. Presented at the Conference for International Cooperation and Partnership, Salzburg-Vienna. 1-7 July, 1962.

RASHTRAPATI BHAVAN,
NEW DELHI-4.
राष्ट्रपति भवन,
नई दिल्ली-4।

July 11, 1963

I am happy to know that Professor P.C.Mahalanobis has recently completed his seventieth year. His outstanding contributions in the field of Statistics have earned him world-wide recognition. His work has been of much use to us in our development planning. I send my best wishes to him for many more years of useful work.

(S.Radhakrishnan)

प्रधान मंत्री भवन
PRIME MINISTER'S HOUSE
NEW DELHI

Dehra Dun,
June 2, 1963

I am happy to join in the felicitations offered to Professor P.C. Mahalanobis on his seventieth birthday. I have known him for several decades and have come into close contact with him particularly since we attained independence and he has been connected with the Central Statistical Department and the Planning Commission. An eminent scientist and the builder up of the Indian Statistical Institute which has become one of the leading centres of statistical work in the world, I have been deeply struck by his broad and comprehensive approach to national development and his astonishing energy. He is full of ideas and it is always a pleasure to discuss any subject with him. The record of his achievements and the work of national development and international cooperation is a formidable one. In paying my tribute of admiration for the work he has done, I earnestly hope that he will have many years more of service to the causes he has devoted himself to.

Jawaharlal Nehru

Academy of Sciences, Bulgaria

May I be allowed in the first place to congratulate you most cordially on the glorious 70th anniversary of your birth and on your long and remarkable scientific and public activities by which you won renown and respect in all countries, including my Homeland, the People's Republic of Bulgaria.

My feelings of joy on the occasion of your 70th birthday are naturally linked with the reminiscences, pleasant to every Bulgarian, of your collaboration with the outstanding Bengalee poet and thinker Rabindranath Tagore whom you accompanied during his tour of Europe in the twenties of this century and with whom you visited our country for the first time.

Bulgarian scientists are well aware of your merits in the field of statistics and of its application in science and practice, as well as of your vast learning and interests in science and culture, and your merits as an outstanding champion of peace and well-being for the peoples.

T. D. Pavlov
Honorary President

Academy of Sciences, Czechoslovakia

Information has been received that the Czechoslovak Academy of Sciences has decided to award the Gold Medal of the Academy for meritorious service to science and mankind to Professor P. C. Mahalanobis.

Academy of Sciences, Poland

On behalf of the Presidium of the Polish Academy of Sciences and the Polish research workers I have the pleasure to extend to you our cordial congratulations on your Seventieth Birthday and to convey you an acknowledgement of your untiring efforts and magnificent successes you have achieved in many fields of learning.

The Polish research workers know you as an eminent scholar, excellent organiser of scientific, economic and social life, and highly appreciate your works and various activities as well as your contribution to India's planned economical development. Your scientific works on economical and statistical topics,—the fields in which you are renowned authority,—have been particularly studied and estimated in our scientific quarters.

May I also add our best congratulations for you as a founder, organisor and Director of the Indian Statistical Institute, now one of the most outstanding world research centres in the field of statistics with which the Polish research workers have already had an opportunity to establish scientific relations.

I should like to send you on this occasion our best wishes for many happy returns and many years in which you will have the opportunity to carry out your ideas. Looking forward to your enjoying a perfect health that will enable you to devote yourself to some new responsible tasks that still await you in the years to come, I remain,

Janusz Groszkowski
President

Academy of Sciences, U.S.S.R.

The Academy of Sciences of U.S.S.R. warmly congratulates you, a scholar of international reputation, one of the prominent figures of Indian Science, on the occasion of your 70th birthday and wholeheartedly wishes you a healthy and long life and further fruitful work in the interest of the Indian people and to the profit of world science.

The Soviet scholars are well aware of your great and multifarious scientific and scientific-pedagogical activities which started nearly half a century ago. We have a high esteem for your distinguished contributions to the development of Indian and world-science. The field of your scientific interests is wide: Physics and Mathematics, Meteorology and Hydrology, Anthropology and Sociology, Demography and Economics. Your achievements in the fields of theoretical and applied statistics have been recognised by the scientific circles throughout the world.

Under your directorship, the Indian Statistical Institute has acquired a well merited authority in the whole world. The investigations conducted in the Institute and its departments and the creation of highly qualified specialists have a great significance for India and also for other countries.

In our country, your name is well known not only to the scholars but also to a wide circle of Soviet public. You are known here as a champion for the cause of progress and development in India, as a fighter for the cause of economic stability in your country.

Dear Professor ! Soviet scholars have a high esteem for your scientific and social achievements, your accomplishments in international scientific collaboration, and for the scientific bonds you have established and strengthened between the Soviet Republic and India.

The recognition of your achievements in world-science is reflected in your having been elected as a foreign member of the Academy of Sciences of U.S.S.R. as a fellow of Royal Society and of many other foreign scientific organizations.

On this Jubilee day, we extend our greetings to you, a citizen of great India, a progressive social worker and a scholar of international reputation.

For and on behalf of the Presidium of the Academy of Sciences, U.S.S.R. and Soviet scholars.

M. V. Keldish
President

American Statistical Association

It is a pleasure to extend to Professor Mahalanobis the congratulations of the American Statistical Association on this, his seventieth birthday. His record of sustained accomplishment in the field of survey methods and sampling theory is too well known to need recounting here. We of the American Statistical Association are particularly indebted to him for his many demonstrations of the value of scientific survey design in government statistical work. His challenging use of interpenetrating samples for the evaluation and control of error in large scale surveys will in the years to come continue to ensure improvement in the field to which he has given so many years of dedicated service. We are especially proud that he is a Fellow of the Association.

Raymond T. Bowman
President

Biometric Society

On the occasion of his 70th birthday, the Biometric Society is most happy to send its greetings to Professor P. C. Mahalanobis. From the foundation of our Society in 1947, Professor Mahalanobis has been a valued member; he served as one of our first Vice Presidents and was for many years an elected member of our Council.

The high repute of Indian work in statistics and biometry owes very much to Professor Mahalanobis's enthusiasm and hard work. By comparison, the Biometric Society is a very new growth. Yet we remember gratefully how, in December 1951, his invitation enabled us to hold our First International Biometric Symposium in the new buildings of the Indian Statistical Institute. The theme of that Symposium was "Biometric Problems in the Prediction and Estimation of the Growth of Plants in Tropical and Subtropical Regions," and Professor Mahalanobis contributed substantially to the proceedings with his paper "On the Size of Sample Cuts and Crop Cutting Experiments in the Indian Statistical Institute : 1939-1950." He also arranged for the publication of the 116-page Proceedings of the Symposium by the Press of the Institute.

In its constitution the Biometric Society defines its objectives as "the advancement of biological science through the development of quantitative theories and the application, development, and dissemination of effective mathematical and statistical techniques." These objectives have been shared so fully by the Indian Statistical Institute under Professor Mahalanobis's direction, that we have always had a strong feeling of kinship with the Institute and its staff as a major center for advancing our aims in this part of the world. Moreover, the continuous flow of visiting statisticians and biometricians from all parts of the world through the doors of the Institute has been an important aid to international cooperation among scientists in all fields that lie close to biometry. The Biometric Society has been particularly eager to encourage this kind of cooperation.

Many officers and members of our Society have enjoyed the hospitality of the Institute, and the personal hospitality offered so graciously by Professor and Mrs. Mahalanobis in their home. I know that I express the hope of all members when I wish Professor Mahalanobis many more years of effective scientific work and vigorous leadership in the subjects to which he has already contributed so impressively.

C. I. Bliss
President

Ecole des Hautes Etudes, Sorbonne, Paris

We express our good wishes and admiration for you on the occasion of your 70th anniversary.

Ferrand Brandel
President of the Section of
Economic and Social Sciences

Geographical Society of the U.S.S.R.

The Geographical Society of the U.S.S.R. under the Academy of Sciences felicitates you, an outstanding scientist, a public and social worker of our day, on the occasion of your 70th birthday.

Your indefatigable efforts for all-sided development of advanced science in your mother country, your constant efforts for widening international cooperation in the interests of peace in the whole world, for progressive humanity and for a rise in the economic and cultural level for the less fortunate regions of the world, for your long-drawn out struggle against colonialism and its heavy legacy in your country as well as in the world—all this has stimulated a feeling of deep respect and friendship for you among the Soviet scientific community as a whole.

One of the illuminative expression of such feelings and recognition of your great scientific service is your unanimous election as a foreign member of the Academy of Sciences, U.S.S.R.

The Soviet geographers are very highly impressed by the nature of your scientific activity, the extent of your vision, your lively interest in natural sciences and researches in the domain of social sciences and your efforts towards establishing intimate creative relation between theory and practice.

Already at the beginning of your scientific career as the Chief meteorologist of Calcutta and as the Head of the Department of Physics of the Presidency College you conducted original scientific investigations to fight the menace of floods in Eastern India and formulated advanced ideas of an integrated approach to the reconstruction of the Indian river system. Even as far back as that, your interest in statistical investigation came to be revealed.

The tasks of building up a new independent India attracted your attention to the problem of scientific planning, to the elaboration of a programme of economic development for the young Indian Republic, to the problems of exploitation of the rich and varied natural resources and rational geographical distribution of the developing productive forces of the country. Well-known are your contributions to the formulation of the scientific principles and elaboration of the five-year plans for economic and cultural construction in India. You can be rightfully considered to be the initiator and champion of intensive studies in the problems of planning and regional planning in the Indian Republic.

The Soviet geographers have the satisfaction that in this important scientific and governmental activity you with your characteristic perspicacity have unceasingly encouraged extensive geographical investigations of the Nature, Economy and Population in India and widening in all possible ways the contacts between the Indian and Soviet scientists and specialists working in the domain of geography as well as in other domains of knowledge.

We respond to your 70th birthday with full assurance of our cooperation and trust in this that the friendship between the peoples of India and U.S.S.R. will go on being strengthened and extended.

S. V. Kalesnik
Vice-President

Indian Mathematical Society

On the occasion of completing his seventieth year, the Indian Mathematical Society conveys its warm felicitations to Professor P. C. Mahalanobis, F.R.S., and wishes him many more years of happy life and service to the country. Starting his career as a Professor of Physics, Professor Mahalanobis took Statistics as a hobby, and steadily built up a school of Statistics, practically from scratch. His name will be ever remembered for his pioneering work in India in the field of Statistics, and for building up a strong and well-planned edifice for it. Our warm congratulations.

C. N. Srinivasiengar
President

Institute of Asia Economic Affairs, Tokyo

On the 70th Anniversary of your birthday I have great pleasure in extending to you my warm congratulation and best wishes for your happiness.

Seiichi Tobata
President

Indian Science Congress Association

In honouring Professor Mahalanobis on his seventieth birthday we are honouring not only a great man and leader of Science, but also one whose contribution in recent years in the field of planning and development of our country has been of the utmost importance: He has been a pioneer in a real sense. He has been one of the main architects of the Indian Science Congress. He was the General President in 1950. His work and example have inspired many, and would continue to be of unfailing source of inspiration for years to come. We pay our respectful homage and tribute to him, and sincerely wish him many years of happy and fruitful life, devoted to the service of science and the betterment of man.

D. S. Kothari
President

Institut de France Academie des Sciences

In collaborating in the collection of memoirs which are going to be dedicated in homage to you I have given expression to the high esteem in which I hold your scientific works and their applications and the advancements which you have made in the education and application of statistics in India.

I can assure you without going in for votes (and I have been authorised to inform you) that the sentiments which I have expressed are also shared by my co-workers in the Academy of Sciences. We also hope that the bonds of friendship which have been forged between your great country and ours during the past years may continue being strengthened. Permit me, my dear colleague, to express my most cordial and best regards for you.

Maurice Frechet

International Statistical Institute

Professor Mahalanobis includes—among the many up-to-date and creative facets of his personality—the secret of remaining dynamic. Only the wisest people of our day know how to keep young—to widen the gap between their actual and biological age. If he were not so well known with his name and biographical details in hundreds of publications, no one would believe that Prasanta Chandra Mahalanobis, with his commanding presence, black hair, lively expression and readiness to act, had reached the seventy year mark, which all over the world, usually means retirement for university professors.

But appearances are belied by a vast output of work which must, of necessity, extend over many years of study and covers, with real enthusiasm and thoroughness, many fields of theoretical and applied statistics—methodological statistics, anthropometry and biometry, agricultural experimentation, meteorology and hydrology, business statistics, operational analysis and many more which it would be hard to catalogue. I have counted 149 titles of his publications between 1922 and 1961—but another two busy years have now gone by and, in any case, we all know how impossible it is to list accurately the whole of any writer's work, as the author himself often finds it a difficult job.

The salient fact which emerges from scanning the list of subject headings so far classified is that here is a mind which is not only open to every branch of scientific research but which acknowledges the need for rigorous method and the practical value of the results achieved, with a feeling for figures and possibilities of error, and is at the same time concerned with finding new device for investigation.

His greatest work has been in the wide field of experimental estimates in agriculture—i.e. a subject which is of vital interest for developing countries like India. But the variety of other important lines of research already mentioned should be noted. It must be added, to complete the picture, that when Professor Mahalanobis takes up a new field of study he dedicates the whole of his efforts to making a thorough investigation of it. He is no dilettant flitting from one thing to another, achieving nothing.

At least two of his methodological contributions must be recorded. One is the introduction of his measurement of generalized distance in multivaried analyses, the achievement of the solution of an old problem, probably encouraged above all by his studies on comparative anthropometry. The other is the development of, to some extent sequential, procedures for carrying out widespread sampling. This innovation is part of a series of improvements introduced in the course of important experimental and census research for coping with some of the fundamental problems of Indian life and economic development. If the country is now well equipped with men and means for carrying out every kind of sample surveying, thus helping enormously to pinpoint and settle many economic and social problems, this is largely due to Professor Mahalanobis and his team of co-workers and followers.

His tireless efforts on behalf of science and his country have led to a number of practical achievements. The first was the creation of the Indian Statistical Institute (1931) which, together with *Sankhyā : The Indian Journal of Statistics*, started shortly afterwards (in 1933), formed the nucleus of a research movement bringing the Indian school of statistics into the top line. Mahalanobis's work has not, however, been confined to India but has won general recognition in international circles. What he has written about the teaching of Statistics in various types of school is also important and so is his educational work at the International Statistical Institute.

His was the great merit of founding the International Statistical Education Centre in Calcutta in 1950, under the auspices of U.N.E.S.C.O. and the Indian

Government, which works in conjunction with the I.S.I. and the Indian Statistical Institute. This centre continues to meet its original aims of giving courses for

(A) teachers of statistics,

(B) officials without experience in statistical work.

By the end of 1961, 378 pupils had been trained at the centre.

As an outstanding Indian personality and as chairman of the Committee of Statistical Education of the Institute, Professor Mahalanobis has given a great boost to statistical education for many years. While he was chairman, other centres—at Beirut and for one year at Rome—were set up and worked well.

This is not the place to list the numerous acknowledgements of Professor Mahalanobis' statistical work in India and elsewhere. His biography appears in *A History of the International Statistical Institute* by J. W. Nixon and the *International Who's Who* contains a list of the distinctions he has won. We cannot, however, fail to mention the tribute paid him by one of the oldest, and certainly the most famous, scientific societies in the world—by making him a Fellow of the Royal Society, London.

The writer of this article is proud to have shared with Professor Mahalanobis the friendship of Sir Ronald Fisher. Now that we are celebrating the Jubilee of the former, it is fitting to mention the together, recalling the memory of the latter who has so recently departed from among us. Professor Mahalanobis and Sir Ronald were together elected honorary presidents of the International Statistical Institute at Stockholm in 1957—a rare, if not unique, distinction which was intended as a solemn mark of our august assembly's appreciation of two such highly qualified men.

I, as President of the Institute, would like to congratulate the honorary president, Professor Mahalanobis, and add my tribute to those pouring in on the occasion of his Jubilee. I would also like to express to him, and to Mrs. Mahalanobis who has been a valuable and constant companion in her husband's work and undertakings, as well as to their great country, the hope for a long life of useful activity by a great mind which, fortunately for us, is still going to achieve many things in the future.

M. Boldrini
President

National Institute of Sciences of India

On behalf of the National Institute of Sciences I have pleasure in sending our felicitations and good wishes to Professor Mahalanobis on his seventieth birthday. Professor Mahalanobis has established the importance of Statistics in India in Government and National Planning. Through his achievements he has enhanced the prestige of Indian Science in other countries. We send him our best wishes for many years of continued fruitful activity.

Homi Bhabha
President

Patrice Lumumba University of Friendship among Peoples

The great scientist and social worker of India, Professor P. C. Mahalanobis completed his 70th birthday on 29 June. His name is connected with the establishment and guidance of the many-sided activity of the Indian Statistical Institute which is engaged in studying problems of physics and mathematics, meteorology and hydrology, anthropology and sociology, demography and economics. The contributions of Professor P. C. Mahalanobis in the domain of theoretical and applied statistics are of outstanding significance.

The researches and practical operations conducted in the Institute and its different departments have not only theoretical scientific interest but also great practical importance for the building up of an independent India. Organisation of teaching courses in the Institute and later of the International Statistical Education Centre have ensured the training of highly qualified cadres of statisticians not only for India but also for several other countries in the East.

From 1949 Professor Mahalanobis has been an honorary Adviser to the Cabinet of Ministers in statistics and from 1955 he has been a member of the Planning Commission of India holding the rank of a minister.

A person of exceptional energy, initiative, outstanding organising talent, a blazing patriot of his mother country, Professor Mahalanobis has made great contributions to the establishment and expansion of the economic and scientific relations between India and the Soviet Union. He has visited the U.S.S.R. a number of times and has lectured several times in the Academy of Sciences, U.S.S.R. of which he happens to be a foreign member, in the Moscow State University, the Patrice Lumumba University of Friendship among Peoples and in other educational institutions of the Soviet Union.

On the occasion of his 70th birthday his numerous Soviet friends and colleagues wish him sound health and long life for the progress of his mother country and further strengthening the relations between India and the Soviet Union and for peace in the whole world.

P. Erzin
Director

The Royal Society, London

The Royal Society of London joins in sending greetings to Professor Prasanta Chandra Mahalanobis, F.R.S., on the occasion of his seventieth birthday.

Elected a Fellow of the Royal Society in 1945, Professor Mahalanobis is distinguished for his outstanding contributions to statistical science.

Professor Mahalanobis may look back with pride on a long and distinguished career. His colleagues in the Fellowship of the Royal Society wish him many more years of fruitful and happy activity.

H. W. Florey
President

Royal Statistical Society

The Indian Statistical Institute has attained a remarkable reputation throughout the world, both for the depth of the theoretical research accomplished there, and the pertinent relation to current world problems of its applied statistical work. This arose almost entirely out of the initial energy and drive of one man—Prasanta Chandra Mahalanobis—whose seventieth birthday this volume commemorates.

In the fields of mathematics, physics, anthropology, meteorology, agronomy, economics and statecraft Professor Mahalanobis has played a distinguished part. That he has been a Fellow of the Royal Society since 1945 and a Foreign Member of the U.S.S.R. Academy of Sciences since 1958 bears witness to his scientific versatility. We ourselves are proud to number him among our Honorary Fellows.

The Royal Statistical Society congratulates Professor Mahalanobis and India on his seventieth birthday and wishes him many more years of useful activity.

J. O. Irwin
President

Dele